Geometric function theory in metric spaces

by

Behnam Esmayli

BSc, University of Tabriz, 2009

MSc, Sharif University of Technology, 2011

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Behnam Esmayli

It was defended on

August 26, 2021

and approved by

Piotr Hajłasz, Department of Mathematics, University of Pittsburgh

Jason DeBlois, Department of Mathematics, University of Pittsburgh

Juan Manfredi, Department of Mathematics, University of Pittsburgh

Paweł Goldstein, Faculty of Mathematics, Informatics and Mechanics, University of Warsaw

Dissertation Director: Piotr Hajłasz, Department of Mathematics, University of Pittsburgh

Geometric function theory in metric spaces

Behnam Esmayli, PhD

University of Pittsburgh, 2021

First, we generalize the coarea inequality, also known as Eilenberg's inequality, and provide a self-contained proof of it. The only previously known proof is based on a difficult result of Davies, which our proof avoids. Next, we find several equivalent conditions for Lipschitz functions from Euclidean cubes into arbitrary metric spaces to have a Lipschitz factorization through a metric tree. As an application we prove a recent conjecture of David and Schul [8]. The techniques developed for the proof of the factorization result yield several other new and seemingly unrelated results. We prove that if f is a Lipschitz mapping from an open set in $\mathbb{R}^n$ onto a metric space X, then the topological dimension of X equals n if and only if X has positive n-dimensional Hausdorff measure. We also prove an area formula for length-preserving maps between metric spaces, which gives, as a concrete application, a new formula for integration on countably rectifiable sets in the Heisenberg groups.

keywords: Hausdorff measure, weighted integrals, coarea inequality, metric derivative, area formula, coarea formula, mapping content, length preserving maps, Heisenberg groups, topological dimension, metric trees, factorization, quasiconvex metric spaces.

Table of contents

Acknowledgement

I would like to thank my advisor who has been an inspiration mathematically and otherwise. I thank my defense committee for their time and guidance.

I would like to thank Prof. Jason Deblois, Prof. Paweł Goldstein, Dr. Tom Everest, and Dr. Daniel Hockensmith for the tens of recommendation letters they sent during a tough job market.

I thank the whole math community at Pitt from faculty to staff to students (graduate and undergraduate) for a supportive and welcoming environment.

I owe special gratitude to my great friend and mentor Mohsen Aliabadi, without whose support and encouragement none of this would have happened.

I would like to thank my family back home for their incessant love and support.

I dedicate this thesis to my wonderful wife who published more babies during my PhD than I did papers!

1.0 Introduction and overview

This thesis is made up of two major parts. In Chapters 3 and 4 we follow [19] and present a self-contained and complete proof of the famed coarea inequality, known also as Eilenberg's inequality. The coarea inequality gives a upper bound on the average size of the fibers of a Lipschitz map between arbitrary metric spaces in terms of the size of the domain and the Lipschitz constant. Being of such broad generality, in this context we cannot bring in many analytic tools. But this changes when we assume our domain to be a Euclidean space.

A great deal of analysis has been developed for the study of Lipschitz maps from Euclidean spaces into arbitrary metric spaces since 90's, beginning with work of Kirchheim [34]. Kirchheim proved an analogue of Rademacher's differentiability theorem for Lipschitz maps into metric spaces by replacing the Fréchet differentiability with the weaker notion of *metric differentiability*. This notion of derivative has proven to be sufficient enough for much of the geometric measure theory theorems, such as the area and coarea formulas, to still hold.

In chapter 5 we give a quick overview of the metric differentiability, including a complete proof of the Kirchheim-Rademacher differentiability theorem. The proofs were known, but we derive it from and emphasize the so-called *componentwise derivative* of maps into ℓ^∞, the Banach space of bounded real sequences. This point of view is essential for many of the later applications and proofs in this thesis.

Chapter 6 we prove a number of equivalent conditions for a Lipschitz map from a Euclidean space into a metric space to factor through a tree. This in particular proves a recent conjecture of David and Schul [8]. This chapter also contains a few other results of independent interest.

After this quick summary, let us now further elaborate on each main chapter individually.

1.1 Chapter 3 summary

Chapter 3 is a prerequisite for the proof of the coarea inequality in Chapter 4. But Chapter 3 is a self-contained and detailed presentation of the theory of weighted integrals, which are of independent interest.

It is easy to imagine that sometimes one may wish to integrate functions that may not be measurable. The upper integral is one solution. If μ is a measure on the space X, then for any function $f : X \to [0, \infty]$, defined μ-a.e. on X, we define its *upper integral* as

$$\int_X^* f \, d\mu = \inf \int_X \phi \, d\mu \,,$$

where the infimum is taken over all *μ-measurable* functions ϕ satisfying $0 \leq f(x) \leq \phi(x)$ for μ-a.e. $x \in X$.

To emphasize, we do note require f to be measurable. Clearly, for measurable functions the upper integral does coincide with the usual integral – take $\phi = f$.

However, this definition still does require knowing/finding a whole class of *measurable* functions on the space. It turns out that when X is a metric space and μ is a Hausdorff measure there is a characterization of the upper integrals that avoids measurability issues altogether.

In Chapter 3 we begin by a few simple Lemmata about upper integrals. Then we recall the definition of the *weighted integrals*, introduced by Federer in his proof of the coarea inequality. Weighted integrals are defined using only metric notions.

The main and quite surprising fact is that the weighted integral equals upper integral (Theorem 45). Federer proved this fact under further assumptions, either on the metric space, or on the support of the function being integrated. But he conjectured that the assumptions are superfluous. Later, Davies proved a highly nontrivial result about the Hausdorff contents and claimed that Federer's conjecture would follow from that. The proof written as a whole can be found in [45].

Our contribution is a new proof of the equality of the weighted integral and upper integral. Our proof is based on an argument that we learned from Nazarov [38]. It is elementary and uses tools only from the metric geometry. In particular, it completely avoids the use of

Davies' result. Our impression is that both the technique used in the proof (Theorem 47) and the notion of weighted integral will find new applications.

Material in Chapter 3 (and Chapter 4) are from our paper [19].

1.2 Chapter 4 summary

Chapter 4 provides an elementary and self-contained proof of the following result which is known under the name of the *coarea inequality* or the *Eilenberg inequality.*

Theorem 1. *Let X and Y be arbitrary metric spaces, $0 \leq t \leq s < \infty$ (any) real numbers and $E \subset X$ any subset. Then, for any Lipschitz map $f : X \to Y$ we have*

$$\int_Y^* \mathcal{H}^{s-t}\left(f^{-1}(y) \cap E\right) \, d\mathcal{H}^t(y) \leq (\mathrm{Lip}\, f)^t \, \frac{\omega_{s-t}\omega_t}{\omega_s} \mathcal{H}^s(E) \,. \tag{1.1}$$

The inequality was first proved by Eilenberg [14] in 1938 in the case when $t = 1$, $Y = \mathbb{R}$ and $f(\cdot) = d(\cdot, x_o) \to \mathbb{R}$ is the distance to a point on a metric space X. Then it was generalized in [15] to the case of $t = 1$, $Y = \mathbb{R}$ and $f : X \to \mathbb{R}$ any Lipschitz function.

It seems however, that a related argument was used by Szpilrajn[1] [13] in the proof that if $\mathcal{H}^{n+1}(X) = 0$, then the topological dimension of X is at most n. Szpilrajn's proof is reproduced in [30, Theorem 7.3] and [28, Theorem 8.15]. Szpilrajn mentions that his argument is based on Nöbeling's proof of a weaker result that the topological dimension is bounded from above by the Hausdorff dimension of a metric space [43] (Nöbeling's paper is reproduced in [42]). The reader may find a translation of Nöbeling's paper in MathOverflow [39], and it is clear that his argument was closely related to Eilenberg's inequality for the distance function. From reading Szpilrajn's paper, it is also clear that there was a strong collaboration between him and Eilenberg.

Remark 2. Most of the proofs that the reader may find in the literature [5, Theorem 13.3.1], [35, Lemma 5.2.4], [41, Theorem 7.7], apply to the case of Lipschitz mappings $f : X \to \mathbb{R}^m$ and $t = m$, and the proofs do not differ much from that in [15]. Since the proofs use the

[1]He changed his name to Marczewski while hiding from Nazi persecution.

fact that for a subset $A \subset Y = \mathbb{R}^m$, the isodiamteric inequality holds, that is $\mathcal{H}^m(A) \leq \omega_m(\operatorname{diam} A)^m/2^m$, there is no obvious way how such proofs could be generalized to other metric spaces Y.

Remark 3. Regarding coarea inequality for mappings into metric spaces one should mention an interesting paper by Malý [37]. The result given in [2, Proposition 3.1.5] covers the general case but, as confirmed by the authors, the proof is incorrect.

Proving the result in a more general case was a remarkable achievement of Federer [21], see also [22, Theorem 2.10.25]. However, he could prove Theorem 1 only under additional assumptions that

(a) The integrand $\mathcal{H}^{s-t}(f^{-1}(y) \cap E)$ is positive (only) on a set of σ-finite measure $\mathcal{H}^t$; or

(b) The space Y is *boundedly compact*, meaning that bounded and closed sets are compact.

His strategy was as follows. He first proved an inequality more or less equivalent to

$$\int_Y^\bullet \mathcal{H}^{s-t}_\delta \left(f^{-1}(y) \cap E\right)\, d\mathcal{H}^t \leq (\operatorname{Lip} f)^t\, \frac{\omega_{s-t}\omega_t}{\omega_s} \mathcal{H}^s(E)\,, \tag{1.2}$$

where the left-hand side is the *weighted integral* we mentioned above (see Definition 38 for the rigorous definition.). Federer [22, 2.10.24] used however, different notation (see Remark 42).

This inequality follows from a straightforward covering argument. In fact the proof is very similar to the classical proof due to Eilenberg, the one the reader can find in [5, 35, 41], see Remark 2.

The coarea inequality then follows from the Theorem 45, namely that

$$\int_Y^* g(y)\; d\mathcal{H}^t(y) = \int_Y^\bullet g(y)\; d\mathcal{H}^t\,. \tag{1.3}$$

and a simple monotone convergence theorem for upper integrals as $\delta \to 0^+$.

Federer [22, 2.10.24] proved this equality result under the restrictive assumption that one of the following two conditions is satisfied: (a') The function g is positive on a set of σ-finite measure $\mathcal{H}^t$; or (b') the space Y is boundedly compact. Therefore he could only prove Theorem 1 under the assumptions (a) or (b) listed above.

While the inequality

$$\int_Y^* g(y)\; d\mathcal{H}^t(y) \geq \int_Y^\bullet g(y)\, d\mathcal{H}^t(y)$$

is easy to prove in the general case (see (3.14)), the problem is to prove the opposite inequality (Federer proved it when (a') or (b') holds true). In the general case, Federer [22, p. 187] stated the following:

The general problem whether or not the preceding inequality can always be replaced by the corresponding equation is unsolved.

The problem was answered in the positive by Davies [10, page 236]:

Note added 8 September 1969. H. Federer tells me that this work answers a question he raised in Geometric measure theory (Berlin, 1969) [...]

There is no explicit proof of (1.3) in the work of Davies, but the main result of Davies [10, Theorem 8, Example 1], provides a missing step in generalizing Federer's proof. In fact it is the celebrated Increasing Sets Lemma [10, Theorem 8] that was needed to complete Federer's proof:

Theorem 4. *Suppose (X, d) is an arbitrary metric space, $t \in [0, \infty)$, and $\delta > 0$. Then for any increasing sequence of subsets $A_1 \subset A_2 \subset A_3 \subset \cdots$,*

$$\mathcal{H}^t_\delta\Big(\bigcup_i A_i\Big) = \lim_{i\to\infty} \mathcal{H}^t_\delta(A_i)\,.$$

With (1.3) being true for an arbitrary metric space Y, Federer's proof of Theorem 1 applies to the case of arbitrary metric spaces X and Y.

From what we could dig out from the literature, it would be fair to call Theorem 1 the Nöbeling-Szpilrajn-Eilenberg-Federer-Davies inequality.

Surprisingly, it wasn't until 2009 when Reichel [45] in his PhD thesis, re-wrote a complete proof of Theorem 1 in its full generality, by following the original proof of Federer while making use of Davies' result. Reichel's thesis seems to be the only place with a complete proof of Theorem 1, except that Reichel did not include the proof of Davies' theorem.

Davies' theorem [10, Theorem 8] (Theorem 4 above) is very difficult and its proof makes use of Ramsey's theorem, ordinal numbers and non-principal ultrafilters.

The proof of the coarea inequality looks very short and simple, but this is thanks to the fact that (1.3) is proved independently in Chapter 3 (using an elementary argument that completely avoids the use of Davies' result).

The rest of the Chapter 4 deals with generalizations of the coarea inequality.

Most of the older applications of Theorem 1 are in the case of Lipschitz mappings $f : X \to \mathbb{R}^m$ and $t = m$. However, in a recent development of analysis on metric spaces, the general version of Theorem 1 plays an increasingly important role. It is a fundamental result and it deserves to have a proof that is self-contained and easy to read. Our proof of how to conclude Theorem 1 from Theorem 45, follows Federer's argument, but we believe is much easier to read than Federer's proof. In writing this proof we also used a presentation of Federer's proof given in [45].

Material in Chapter 3 and Chapter 4 comprise our paper [19].

Acknowledgement. We would like to thank Mikhail Korobkov for discussions on topics related to Definition 54.

1.3 Chapter 5 summary

Chapters 5 and 6 are concerned with Lipschitz functions from Euclidean spaces to arbitrary metric spaces. The additional structure on the domain allows many of the analytical tools to be applicable. The starting point is a notion of differentiability for such maps introduced by Kirchheim in [34].

Chapter 5 is an overview of the theory of *metric differentiability*, including a proof of the Kirchheim-Rademacher theorem which claims that Lipschitz maps from Euclidean domains into arbitrary metric spaces are metrically differentiable almost everywhere.

We also emphasize the related notion of componentwise derivatives for maps into ℓ^∞, the space of bounded real sequences. The insistence of this particular target space is not restrictive since every separable metric space has an isometric embedding in ℓ^∞, by the famed Kuratowski-Fréchet embedding theorem.

The notion of metric derivative is strong enough to give area and coarea formulas for maps into metric spaces. We do mention them toward the end of Chapter 5. However, our use of metric derivatives will be mainly through their connection to the *mapping contents* in Chapter 6 and the result about factorization of maps through metric trees. A few Lemmata we prove in Chapter 5 will be needed in Chapter 6.

1.4 Chapter 6 summary

The first main result in Chapter 6 is as follows (Please see Theorem 108 for the full statement.)

Theorem 5. *If $f : [0,1]^{n+m} \to X$, $n \geq 1$, $m \geq 0$, is a Lipschitz map into a metric space, and $E \subset [0,1]^{n+m}$ is a measurable set, then the following conditions are equivalent:*

(A) $\operatorname{rank}\operatorname{md}(f,x) \leq n-1$ *a.e.*

(B) $\mathcal{H}^{n,m}_{\infty}(f,E) = 0$.

(C) $\Theta^{*n}(f,x) = 0$ *a.e.*

Here $\operatorname{md}(f,x)$ stands for the metric derivative of the function f at the point x. We discussed this notion in previous sections. It is a seminorm on $\mathbb{R}^{n+m}$. There is a well-defined notion of rank for seminorms on Euclidean spaces. This explains the notation in (A). Notation in (B) and (C) are explained in the next paragraphs.

This theorem connects two previously known generalizations of the classical implicit function theorems for Lipschitz maps f from (subsets of) the Euclidean space $\mathbb{R}^{n+m}$ into an arbitrary metric space X in the sense that after a C^1 (local) change of coordinates the fibers of the map (i.e. the preimages of singletons) are straightened out to be contained in parallel copies of $\mathbb{R}^m$.

The first generalization was in [4], where the authors introduced the (n,m) *mapping content* $\mathcal{H}^{n,m}_{\infty}(f,E)$ where $E \subset Q_0 = [0,1]^{n+m}$. They showed that if $\mathcal{H}^{n,m}_{\infty}(f,Q_0) > 0$ (along with ither less stringent conditions), then there are subsets of positive measure on which the implicit function theorem holds. Their results were quantitative. In [27] the authors introduce *the upper n-density* $\Theta^{*n}(f,x)$ and then proved an implicit function theorem on the subset where $\Theta^{*n}(f,x) > 0$.

Thus, the theorem above gives a new condition (negation of (A) to be precise) for when the implicit function theorem holds and confirms that the conditions in [4] and [27] are equivalent to saying that “the derivative of the map is full-rank”, which reminds one of the familiar condition in the Euclidean implicit function theorem.

In [8] the authors conjecture [8, Conjecture 1.13] that if $f\colon Q_o = [0,1]^3 \to X$ satisfies

$\mathcal{H}^{2,1}_\infty(f, Q_o) = 0$, then f factors through a metric tree. Recall that a *metric tree*, also known as an $\mathbb{R}$-*tree*, is a geodesic space which contains no subsets homeomorphic to $\mathbb{S}^1$, so it is a geodesic space without "loops". Given metric spaces X, Y, Z, and a Lipschitz map $f : X \to Y$, we say that f *factors through* Z if there are Lipschitz mappings $\psi : X \to Z$ and $\phi : Z \to Y$ such that $f = \phi \circ \psi$.

We prove a slightly more general version of this conjecture:

Theorem 6. *If $f : Q_o = [0,1]^n \to X$, $n \geq 2$, is a Lipschitz map into a metric space, then the following are equivalent:*

(*) *f factors through a metric tree.*

(A') $\operatorname{rank}\operatorname{md}(f, x) \leq 1$ *a.e.*

(B') $\mathcal{H}^{2,n-2}_\infty(f, Q_o) = 0$.

(C') $\Theta^{*2}(f, x) = 0$ *a.e.*

See Theorem 106 for a complete list of the equivalent conditions for factorization through a tree. Equivalence of (A'), (B') and (C') is given by the equivalence of (the negations of) (A), (B) and (C) in Theorem 5. So, the theorem is a consequence of the following.

Theorem 7. *If $f : [0,1]^n \to X$, $n \geq 1$, is a Lipschitz map into a metric space, then f factors through a metric tree if and only if* $\operatorname{rank}\operatorname{md}(f, x) \leq 1$ *almost everywhere.*

The techniques developed for the factorization through metric trees, Theorem 6 and its full version Theorem 106, yield several seemingly unrelated results. We mention two here.

Theorem 8. *Suppose that $f : \mathbb{R}^n \supset \Omega \to X$ is a Lipschitz continuous map from an open set onto a metric space X, $f(\Omega) = X$. Then,* $\dim X = n$ *if and only if* $\mathcal{H}^n(X) > 0$.

Here $\dim X$ stands for the topological dimension defined in section 6.4. This theorem is relabelled and proved as Theorem 134.

The next result (relabelled as Theorem 112 in Chapter 6) gives a new formula for integration on countably rectifiable sets in the Heisenberg groups. Here $\mathbb{H}^n$ is the n'th Heisenberg group modelled on $\mathbb{R}^{2n+1}$ and $\pi \colon \mathbb{H}^n \to \mathbb{R}^{2n+1}$ is the projection onto the fist $2n$-coordinates. We deduce this from a general area formula that we prove for the length preserving maps between metric spaces (Theorem 123).

Theorem 9. *Assume that a set $E \subset \mathbb{H}^n$ is countably k-rectifiable for some $k \leq n$. Then for any Borel function $g : E \to [0, \infty]$, we have*

$$\int_E g(x)\, d\mathcal{H}^k_{cc}(x) = \int_{\pi(E)} \Big(\sum_{x \in \pi^{-1}(y) \cap E} g(x) \Big)\, d\mathcal{H}^k(y).$$

1.5 The layout of the dissertation

Chapter 2 collects basic and standard tools from measure theory and metric spaces. The results are known but in some sections we will provide careful presentation with detailed proofs of the lesser known results.

Chapter 3 introduces the weighted integrals and gives a new and elementary proof that they coincide with the upper integral (Theorem 45).

In Chapter 4 we begin by a history of the coarea inequality. Then we introduce the mapping content $\Phi^{s,t}$ (Definition 54) and prove a generalizations of the coarea inequality (Theorem 58).

The results in Chapters 3 and 4 are accepted for publication as [19].

Chapter 5 includes a detailed proof of the Kirchheim-Rademacher theorem regarding differentiability of Lipschitz maps into metric spaces. We derive this from the componentwise derivative of $f \colon \Omega \to \ell^\infty$ (Theorem 82). Proposition 99 is a new result, which is needed in Chapter 6. The chapter ends with the statements of the Euclidean as well as metric area and coarea formulas.

Chapter 6 uses the techniques of Chapter 5 to prove the main theorem 106 which gives multiple equivalent conditions for when a map factors through a metric tree. But the chapter contains multiple other new results. In Section 6.2 we prove a few lemmata regarding the rank of metric derivative. In Section 6.3 we deduce an area formula for length preserving maps (Theorem 123). It gives a transformation formula for integration over rectifiable sets in the Heisenberg groups. Section 6.4 contains a new result regarding the topological dimension of images of Lipschitz maps defined on open subsets of Euclidean spaces (Theorem 134). In Section 6.5 we discuss a well known and a general construction of a factorization of a Lipschitz map $f : X \to Y$ defined on a quasiconvex metric space. In the final two sections, we use the tools from previous sections to prove the main theorems.

1.6 Notation throughout thesis

In a metric space (X, d), the open and closed balls of radius $r > 0$ centered at x will be denoted by $B(x, r) = \{y : d(x, y) < r\}$ and $\bar{B}(x, r) = \{y : d(x, y) \leq r\}$, respectively. Closure of a set E will be denoted by $\bar{E}$; as a warning, note that in general closed ball might be strictly larger than the closure of the open ball. Symbol B will always be used to denote a ball, open or closed. If $B = B(x, r)$ is a ball, $\sigma B = B(x, \sigma r)$, $\sigma > 0$, will denote a dilated ball (the same notation is used for closed balls).

The characteristic function of a set E will be denoted by χ_E.

A metric space is *boundedly compact* if bounded and closed sets are compact.

A map $f : X \to Y$ between metric spaces is called *Lipschitz* if there exists an $L \geq 0$ such that $d_Y(f(x), f(y)) \leq L d_X(x, y)$ for all x and y in X. The smallest such L, denoted $\operatorname{Lip} f$, is *the Lipschitz constant* of f.

The integral average will be denoted by the barred integral:

$$\fint_E f\, d\mu = \frac{1}{\mu(E)} \int_E f\, d\mu.$$

Hausdorff measure will be denoted by $\mathcal{H}^s$. It is normalized so that on $\mathbb{R}^n$ the measure $\mathcal{H}^n$ coincides with the Lebesgue measure, see Section 2.3 for more detail. We will use $\mathcal{H}^n(E)$, $\mathcal{L}^n(E)$, and $|E|$ to denote the Lebesgue measure of $E \subset \mathbb{R}^n$.

For $A \subset X$, $\operatorname{diam} A = \sup\{d(x, y) : x, y \in A\}$ and

$$\zeta^s(A) = \frac{\omega_s}{2^s} (\operatorname{diam} A)^s, \quad \text{where} \quad \omega_s = \frac{\pi^{s/2}}{\Gamma(\frac{s}{2} + 1)}.$$

Here Γ is *the* gamma function. Note that ω_n is the volume of the unit ball in $\mathbb{R}^n$ so $\zeta^n(B^n(0, r)) = \mathcal{H}^n(B^n(0, r))$. We agree that $\zeta^0(A) = 1$ if $A \neq \varnothing$ and $\zeta^0(\varnothing) = 0$.

For $\delta \in (0, \infty]$, a covering $E \subset \bigcup_{i=1}^{\infty} A_i$ by bounded sets satisfying $\operatorname{diam} A_i \leq \delta$ for all $i \in \mathbb{N}$, is called a *δ-covering of E.* An *open* (*closed*) δ-covering is one where every A_i is open (closed).

The unit ball and the unit sphere (centered at 0) in $\mathbb{R}^n$ will be denoted by $\mathbb{B}^n$ and $\mathbb{S}^{n-1}$.

The (small inductive) topological dimension on X is denoted by $\dim X$.

We write $A \lesssim B$ if there is a constant $C > 0$ that depends only on dimensional data such that $A \leq CB$. If we know that C depends on, say, n and m we shall write $A \lesssim_{n,m} B$ as well.

If μ is a measure, then a property holds *μ-a.e.* (or simply a.e. if μ is understood) if it holds everywhere except for a set of μ-measure zero. Sets with measure zero are also called *null* sets.

If $f\colon X \to Y$, then we use the shorthand $f^{-1}(y)$ in place of $f^{-1}(\{y\})$ to denote the preimage of a single point $y \in Y$. We will never have inverse of functions, so, this is safe.

2.0 Preliminaries

This chapter contains basic tools and well-known results that will be used in the subsequent chapters. Yet, we provide proofs for many of the results for the sake of completeness.

2.1 Upper integral

Definition 10. For a function $f : X \to [0, \infty]$ defined μ-a.e. on X, the *upper integral* is defined by

$$\int_X^* f \, d\mu = \inf \int_X \phi \, d\mu \, ,$$

where the infimum is taken over all μ-*measurable* functions ϕ satisfying $0 \le f(x) \le \phi(x)$ for μ-a.e. $x \in X$.

We do note require f to be measurable. Clearly, for measurable functions the upper integral coincides with the Lebesgue one. Note also that

$$\text{if } \int_X^* f \, d\mu = 0, \text{ then } f = 0 \ \mu\text{-almost everywhere and hence } f \text{ is measurable.} \tag{2.1}$$

Lemma 11. *Let $f_n : X \to [0, \infty]$ a monotone sequence of (not necessarily measurable) functions, i.e. $0 \le f_1(x) \le f_2(x) \le \dots$ for μ-a.e. $x \in X$. If $f(x) := \lim_{n\to\infty} f_n(x)$, then*

$$\lim_{n\to\infty} \int_X^* f_n \, d\mu = \int_X^* f \, d\mu \, . \tag{2.2}$$

Proof. Throughout the proof inequalities between functions are assumed to hold μ-a.e. Clearly the limit on the left hand side of (2.2) exists and

$$\lim_{n\to\infty} \int_X^* f_n \, d\mu \le \int_X^* f \, d\mu. \tag{2.3}$$

Choose measurable functions ϕ_n such that $0 \le f_n \le \phi_n$ and

$$\int_X \phi_n \, d\mu \le \int_X^* f_n \, d\mu + 2^{-n}$$

This and Fatou's lemma yield

$$\int_X^* f\,d\mu = \int_X^* \lim_{n\to\infty} f_n\,d\mu \le \int_X^* \liminf_{n\to\infty} \phi_n\,d\mu \le \liminf_{n\to\infty} \int_X \phi_n\,d\mu \le \lim_{n\to\infty} \int_X^* f_n\,d\mu,$$

which together with (2.3) proves (2.2). $\square$

Definition 12. We say $\phi : X \to [0,\infty]$ is a *step function* if it is μ-measurable and has at most countably many values (we allow infinite values). That is, ϕ is a step function if there exist disjoint μ-measurable subsets $A_i \subset X$ and $0 < a_i \le \infty$ such that

$$\phi(x) = \sum_{i=1}^{\infty} a_i \chi_{A_i}(x)\,; \tag{2.4}$$

We clearly have

$$\int_X \phi\,d\mu = \sum_{i=1}^{\infty} a_i\,\mu(A_i)\,.$$

The following lemma claims that in the definition of the upper integral we may as well restrict to minimizing over step functions.

Lemma 13. *Let $f : X \to [0,\infty]$ any function. Then*

$$\int_X^* f\ \,d\mu = \inf \int_X \phi\,d\mu\,,$$

where the infimum is over all step functions ϕ satisfying $0 \le f(x) \le \phi(x)$ for all $x \in X$.

Proof. Since the claim is true when $\int_X^* f\,d\mu = \infty$, we can assume that $\int_X^* f\,d\mu < \infty$. We can also assume that f is measurable since the general case will easily follow from the definition of the upper integral. For $i \in \mathbb{Z}$ and $1 < \lambda < \infty$ define

$$A_\infty = \{x :\ f(x) = +\infty\}, \quad \text{and} \quad A_i^\lambda = \{x :\ \lambda^i \le f(x) < \lambda^{i+1}\}.$$

Then

$$f \le \phi_\lambda \le \lambda f, \quad \text{where} \quad \phi_\lambda = \infty \cdot \chi_{A_\infty} + \sum_{i\in\mathbb{Z}} \lambda^{i+1} \chi_{A_i^\lambda}$$

and

$$\int_X f\,d\mu \le \int_X \phi_\lambda\,d\mu \le \lambda \int_X f\,d\mu \to \int_X f\,d\mu \quad \text{as } \lambda \to 1^+$$

completes the proof. $\square$

2.2 Covering lemmas

A familiar $5r$-covering lemma, known also as a Vitali type covering lemma, asserts that from any family $\mathcal{F}$ of balls with bounded radii in a metric space, we can select a subfamily $\mathcal{F}'$ of pairwise disjoint balls such that balls in $\mathcal{F}'$ dilated 5 times, cover all balls in $\mathcal{F}$, see e.g. [47, Theorem 3.3]. A close inspection of the proof reveals that we do not really use the fact that this is a family of balls since the proof is based on simple estimates for diameters. Therefore, the lemma holds true for any family of uniformly bounded sets, provided we give a proper meaning of being dilated 5 times. This gives (cf. [22, Section 2.8])

Lemma 14. *Let $\mathcal{F}$ be a family of bounded sets in a metric space such that* $\sup\{\operatorname{diam} F : F \in \mathcal{F}\} < \infty$. *Then, there is a subfamily $\mathcal{F}' \subset \mathcal{F}$ of pairwise disjoint sets such that*

$$\bigcup_{F\in\mathcal{F}} F \subset \bigcup_{F'\in\mathcal{F}'} \mathfrak{5}F',$$

where

$$\mathfrak{5}F' = \bigcup \{F \in \mathcal{F} :\ F \cap F' \neq \varnothing,\ \operatorname{diam} F \leq 2 \operatorname{diam} F'\}.$$

Moreover, if $F \in \mathcal{F}$, then there is $F' \in \mathcal{F}'$ such that $F \cap F' \neq \varnothing$ and $F \subset \mathfrak{5}F'$.

Remark 15. That is $\mathfrak{5}F'$ is the union of F' and all sets that intersect it and have relative small diameter. Clearly $\operatorname{diam} \mathfrak{5}F' \leq 5 \operatorname{diam} F'$.

Proof. Let $\sup\{\operatorname{diam} F :\ F \in \mathcal{F}\} = R < \infty$ and let

$$\mathcal{F}_j = \left\{ F \in \mathcal{F} :\ \frac{R}{2^j} < \operatorname{diam} F \leq \frac{R}{2^{j-1}} \right\}.$$

So, $\bigcup_{j=1}^{\infty} \mathcal{F}_j$ includes all of $\mathcal{F}$ except possibly for some singletons – sets of diameter zero.

We define $\mathcal{F}_1' \subset \mathcal{F}_1$ to be a maximal family of pairwise disjoint sets in $\mathcal{F}_1$. Suppose that the families $\mathcal{F}_1', \ldots, \mathcal{F}_{j-1}'$ have already been defined. Then we define $\mathcal{F}_j'$ to be a maximal family of pairwise disjoint sets in

$$\{F \in \mathcal{F}_j : F \cap F' = \varnothing \text{ for all } F' \in \mathcal{F}_1' \cup \ldots \cup \mathcal{F}_{j-1}'\}.$$

Set $\mathcal{F}' = \bigcup_{j=1}^{\infty} \mathcal{F}_j'$. Every set $F \in \mathcal{F}_j$ intersects with a set $F' \in \bigcup_{i=1}^{j} \mathcal{F}_i'$; it follows that $\operatorname{diam} F \leq 2 \operatorname{diam} F'$ and hence $F \subset \mathfrak{5}F'$.

If there are any singletons $F = \{x\} \in \mathcal{F}$ such that $x \notin \bigcup_{F' \in \mathcal{F}'} F'$, then add F to the collection $\mathcal{F}'$. The updated $\mathcal{F}'$ will remain disjointed and now it satisfies the claim of the lemma. □

Definition 16. Let $\mathcal{F}$ be a family of sets in a metric space X. We say that the family $\mathcal{F}$ is a *fine covering* of a set $A \subset X$ if for every $x \in A$ and every $\varepsilon > 0$, there is $F \in \mathcal{F}$ such that $x \in F \subset B(x, \varepsilon)$.

Corollary 17. *If $\mathcal{F}$ is a family of* closed *sets that forms a fine covering of $A \subset X$, $\sup\{\operatorname{diam} F : F \in \mathcal{F}\} < \infty$, and $\mathcal{F}'$ is as in Lemma 14, then for any finite collection of sets $F_1', \ldots, F_N' \in \mathcal{F}'$ we have*

$$A \subset \bigcup_{j=1}^{N} F_j' \cup \bigcup_{F' \in \mathcal{F}' \setminus \{F_1', \ldots, F_N'\}} \text{ƀ}F' \tag{2.5}$$

Proof. If $x \in A \setminus \bigcup_{j=1}^{N} F_j'$, since the sets F_j' are closed, a ball $B(x, \varepsilon)$ is disjoint with the sets F_j'. If $x \in F \subset B(x, \varepsilon)$, $F \in \mathcal{F}$, then there is $F' \in \mathcal{F}'$ such that $F \cap F' \neq \varnothing$ and $x \in F \subset \text{ƀ}F'$. Since $F \subset B(x, \varepsilon)$ and $B(x, \varepsilon) \cap F_j' = \varnothing$, $F' \neq F_j'$ and hence F' is one of the sets on the right hand side of (2.5).

□

2.3 The Hausdorff measure

Let (X, d) be a metric space. Fix an $0 \leq s < \infty$. For a subset E of X and a $\delta \in (0, \infty]$, the *Hausdorff contents* $\mathcal{H}_\delta^s$ and $\mathscr{H}_\delta^s$ are defined by

$$\mathcal{H}_\delta^s(E) = \inf \sum_{i=1}^{\infty} \zeta^s(A_i), \quad \text{and} \quad \mathscr{H}_\delta^s(E) = \inf \sum_{i=1}^{\infty} \zeta^s(U_i)$$

where the infima are taken, respectively, over all countable coverings $E \subset \bigcup_{i=1}^{\infty} A_i$ by bounded sets with $\operatorname{diam} A_i \leq \delta$ for all $i \in \mathbb{N}$, and over all countable coverings $E \subset \bigcup_{i=1}^{\infty} U_i$ by *open* sets with $\operatorname{diam} U_i \leq \delta$ for all $i \in \mathbb{N}$, in other words, over all δ-coverings and over all *open* δ-coverings. If no such covering(s) exists, we set the corresponding content equal to $+\infty$.

Note that we can always assume that the sets A_i are closed since taking the closure of a set does not increase its diameter. Note also that for any $0 < \varepsilon < \delta < \infty$

$$\mathcal{H}^s_\delta(E) \le \mathscr{H}^s_\delta(E) \le \mathcal{H}^s_{\delta-\varepsilon}(E),$$

because any $(\delta - \varepsilon)$-covering can be enlarged to an open δ-covering with an arbitrarily small increase in diameters of the sets.

The functions $\delta \mapsto \mathcal{H}^s_\delta(E)$ and $\delta \mapsto \mathscr{H}^s_\delta(E)$ are non-increasing, hence for $0 \le s < \infty$

$$\mathcal{H}^s(E) := \lim_{\delta \to 0^+} \mathcal{H}^s_\delta(E) = \sup_{\delta > 0} \mathcal{H}^s_\delta(E) = \lim_{\delta \to 0^+} \mathscr{H}^s_\delta(E) = \sup_{\delta > 0} \mathscr{H}^s_\delta(E)\,,$$

is well-defined. This is the *s-dimensional Hausdorff measure* on X.

Note that $\mathcal{H}^0$ is the *counting measure*, i.e. $\mathcal{H}^0(E)$ equals the number of elements of E.

The Hausdorff measure is an outer measure defined on all subsets of X and all Borel sets are $\mathcal{H}^s$-measurable.

Lemma 18. *$\mathcal{H}^n = \mathcal{H}^n_\infty$ on all subsets of $\mathbb{R}^n$, and $\mathcal{H}^n = \mathcal{H}^n_\infty = \mathcal{L}^n$ on all Lebesgue measurable sets in $\mathbb{R}^n$.*

The following lemma is an easy consequence of the definitions.

Lemma 19. *For any $\alpha \ge 0$ and $E \subset X$, $\mathcal{H}^\alpha_\infty(E) = 0$ if and only if $\mathcal{H}^\alpha(E) = 0$.*

The next result proves that the Hausdorff measure is Borel-regular.

Lemma 20. *For $s \in [0, \infty)$ and every $E \subset X$ there is a decreasing sequence of open sets $V_1 \supset V_2 \supset \ldots \supset E$ such that $E \subset \tilde{E} := \bigcap_{i=1}^\infty V_i$ and $\mathcal{H}^s(E) = \mathcal{H}^s(\tilde{E})$.*

Proof. If $\mathcal{H}^s(E) = \infty$ then we can take $V_i = X$, for all $i \in \mathbb{N}$. So, assume $\mathcal{H}^s(E) < \infty$. For each $i \in \mathbb{N}$ there is a $1/i$-covering $E \subset \bigcup_{j=1}^\infty U_{ij} := U_i$ by open sets, such that

$$\sum_{j=1}^\infty \zeta^s(U_{ij}) \le \mathscr{H}^s_{1/i}(E) + \frac{1}{i} \quad \text{so} \quad \mathscr{H}^s_{1/i}(U_i) \le \mathcal{H}^s(E) + \frac{1}{i}\,.$$

Let $V_i = \bigcap_{k=1}^i U_k$, then $\tilde{E} = \bigcap_{i=1}^\infty U_i = \bigcap_{i=1}^\infty V_i$ has the required properties. $\square$

As an immediate consequence we get

Lemma 21. *If $0 \le s < \infty$, $\mathcal{H}^s(X) < \infty$ and $E \subset X$ is any set, then*

$$\mathcal{H}^s(E) = \inf\{\mathcal{H}^s(U) : U \supset E,\ U \text{ is open}\}. \tag{2.6}$$

The next result is slightly less obvious

Lemma 22. *Let $E \subset X$ be any $\mathcal{H}^s$-measurable set, $0 \le s < \infty$. If $\mathcal{H}^s(E) < \infty$ then*

$$\mathcal{H}^s(E) = \sup\{\mathcal{H}^s(C) : C \subset E,\ C \text{ is closed}\}.$$

Proof. It is enough to prove that for any $\varepsilon > 0$ there exists an F_σ-set contained in E with $\mathcal{H}^s$-measure larger than $\mathcal{H}^s(E) - \varepsilon$.

Fix $\varepsilon > 0$. Let $\tilde{E} = \bigcap_{i=1}^\infty V_i$, $\mathcal{H}^s(\tilde{E}) = \mathcal{H}^s(E)$ be the G_δ set from Lemma 20. Since E is measurable and has finite measure, $\mathcal{H}^s(\tilde{E} \setminus E) = 0$. Each of the open sets V_i is a union of an increasing sequence of closed sets. Since E is contained in that union, there is a closed set $F_i \subset V_i$ such that $\mathcal{H}^s(E \setminus F_i) < \varepsilon/2^i$ and hence the closed set $F = \bigcap_{i=1}^\infty F_i \subset \bigcap_{i=1}^\infty V_i = \tilde{E}$ satisfies

$$\mathcal{H}^s(E \setminus F) = \mathcal{H}^s\Big(\bigcup_{i=1}^\infty (E \setminus F_i)\Big) < \varepsilon.$$

Since $\mathcal{H}^s(F \setminus E) \le \mathcal{H}^s(\tilde{E} \setminus E) = 0$, by Lemma 20, there exits a G_δ-set G such that $F \setminus E \subset G$ and $\mathcal{H}^s(G) = 0$. The proof is complete since

$$\mathcal{H}^s(F \setminus G) = \mathcal{H}^s(F) \ge \mathcal{H}^s(E) - \mathcal{H}^s(E \setminus F) > \mathcal{H}^s(E) - \varepsilon\,,$$

and $F \setminus G$ is an F_σ-set contained in E. $\square$

Lemma 23. *If $s \in [0, \infty)$ and $A_1 \subset A_2 \subset \ldots$ is an increasing sequence of (not necessarily measurable) sets, then*

$$\mathcal{H}^s\Big(\bigcup_{i=1}^\infty A_i\Big) = \lim_{i \to \infty} \mathcal{H}^s(A_i). \tag{2.7}$$

Proof. It suffices to prove that the right hand side of (2.7) is greater than or equal to the left hand side; the opposite inequality is obvious. Let $\hat{A}_i$ be a Borel set such that $A_i \subset \hat{A}_i$, and $\mathcal{H}^s(A_i) = \mathcal{H}^s(\hat{A}_i)$. Let $\tilde{A}_i = \bigcap_{j=i}^\infty \hat{A}_i$. Then $\tilde{A}_i$ is Borel, $A_i \subset \tilde{A}_i$ and $\mathcal{H}^s(A_i) = \mathcal{H}^s(\tilde{A}_i)$. Since $\tilde{A}_1 \subset \tilde{A}_2 \subset \ldots$ are measurable, we have

$$\mathcal{H}^s\Big(\bigcup_{i=1}^\infty A_i\Big) \leq \mathcal{H}^s\Big(\bigcup_{i=1}^\infty \tilde{A}_i\Big) = \lim_{i\to\infty} \mathcal{H}^s(\tilde{A}_i) = \lim_{i\to\infty} \mathcal{H}^s(A_i).$$

□

If a set F is bounded, then $\mathcal{H}^s_\infty(F) \leq \zeta^s(F)$ is an obvious estimate. However, in general we may expect that $\mathcal{H}^s(F)$ is much larger than $\zeta^s(F)$. Indeed, sets with small diameters may have arbitrarily large Hausdorff measure. There is no need to convince the reader that life would be much easier if we could estimate $\mathcal{H}^s(F)$ in terms of the diameter, say $\mathcal{H}^s(F) \leq (1+\varepsilon)\zeta^s(F)$ for some small ε. The next result shows that in fact, in spaces of finite measure, at *almost all* locations and *all* small scales this estimate is true.

Lemma 24. *Let $0 \leq s < \infty$ and $\varepsilon > 0$. If $\mathcal{H}^s(X) < \infty$, then there is a set $E \subset X$ of measure zero, $\mathcal{H}^s(E) = 0$, such that*

$$\forall\, x \in X \setminus E \;\; \exists\, \delta_x > 0 \;\; \forall\, F \subset X \quad \big(x \in F \subset \bar{B}(x, \delta_x) \;\Rightarrow\; \mathcal{H}^s(F) \leq (1+\varepsilon)\zeta^s(F)\big). \quad (2.8)$$

Remark 25. We do not assume measurability of the sets F.

Proof. The claim is obvious for $s = 0$, so assume $s > 0$. Since $\zeta^s(F) = \zeta^s(\bar{F})$, it suffices to prove (2.8) for closed sets F. Let $E \subset X$ be the set of all points $x \in X$ such that for every $j \in \mathbb{N}$, there is a closed set $F_{x,j}$ satisfying

$$x \in F_{x,j} \subset \bar{B}(x, 1/j) \quad \text{and} \quad \mathcal{H}^s(F_{x,j}) > (1+\varepsilon)\zeta^s(F_{x,j}).$$

Clearly, with this definition of E, (2.8) is true and it remains to show that $\mathcal{H}^s(E) = 0$. Suppose to the contrary $\mathcal{H}^s(E) > 0$. According to Lemma 21, there is an open set U such that $E \subset U$ and $\mathcal{H}^s(U) < \mathcal{H}^s(E)(1+\varepsilon/4)$. Given $\delta > 0$, the family

$$\mathcal{F} = \{F_{x,j} : F_{x,j} \subset U,\ j \geq 10/\delta,\ x \in E\}$$

is a fine covering of E by closed sets. Note that $F_{x,j} \subset \bar{B}(x, 1/j)$, $\operatorname{diam} F_{x,j} \leq 2/j \leq \delta/5$. Lemma 14 yields $\mathcal{F}' \subset \mathcal{F}$ such that

$$E \subset \bigcup_{F' \in \mathcal{F}'} \mathfrak{b}F',$$

and the closed sets $F' \in \mathcal{F}'$ are pairwise disjoint. Since $\mathcal{H}^s(X) < \infty$, only countably many of them may have positive measure and the sum of measures is finite so there is a finite collections of sets $F'_1, \ldots, F'_N \in \mathcal{F}'$ such that

$$\sum_{F' \in \mathcal{F}' \setminus \{F'_1, \ldots, F'_N\}} \mathcal{H}^s(F') < 5^{-s} \mathcal{H}^s(E) \frac{\varepsilon}{4}\,.$$

According to Corollary 17,

$$E \subset \bigcup_{j=1}^{N} F'_j \;\cup \bigcup_{F' \in \mathcal{F}' \setminus \{F'_1, \ldots, F'_N\}} \mathfrak{b}F'.$$

Since for each of the sets $F' \in \mathcal{F}'$ we have, $F' \subset U$, $\operatorname{diam} \mathfrak{b}F' \leq 5 \operatorname{diam} F' \leq \delta$,

$$\begin{aligned}
\mathcal{H}^s_\delta(E) &\leq \sum_{j=1}^{N} \zeta^s(F'_j) + \sum_{F' \in \mathcal{F}' \setminus \{F'_1, \ldots, F'_N\}} \zeta^s(\mathfrak{b}F') \\
&\leq \sum_{j=1}^{N} \zeta^s(F'_j) + \sum_{F' \in \mathcal{F}' \setminus \{F'_1, \ldots, F'_N\}} 5^s \zeta^s(F') \\
&\leq \frac{1}{1+\varepsilon} \left(\sum_{j=1}^{N} \mathcal{H}^s(F'_j) + \sum_{F' \in \mathcal{F}' \setminus \{F'_1, \ldots, F'_N\}} 5^s \mathcal{H}^s(F') \right) \\
&\leq \frac{1}{1+\varepsilon} \left(\mathcal{H}^s(U) + \mathcal{H}^s(E) \frac{\varepsilon}{4} \right) \leq \mathcal{H}^s(E) \frac{1+\varepsilon/2}{1+\varepsilon}.
\end{aligned}$$

The estimate is independent of δ so letting $\delta \to 0^+$ we get

$$\mathcal{H}^s(E) \leq \mathcal{H}^s(E) \frac{1+\varepsilon/2}{1+\varepsilon} < \mathcal{H}^s(E)$$

which is a clear contradiction. $\square$

2.4 Rectifiable curves in metric spaces

A *curve* in a metric space (X, d) is a continuous map $\gamma : [a, b] \to X$. The *length* of γ is defined as

$$\ell(\gamma) = \sup \left\{ \sum_{i=0}^{n-1} d(\gamma(t_i), \gamma(t_{i+1})) \right\},$$

where the supremum is over all $n \in \mathbb{N}$ and all partitions $a = t_0 < t_1 < \ldots < t_n = b$. A curve is *rectifiable* if $\ell(\gamma) < \infty$. We will also use notation $\ell_X(\gamma)$. For more information about rectifiable curves in metric spaces, see e.g. [23].

Every rectifiable curve can be reparametrized as a Lipschtz curve [23, Theorem 3.2] so without loss of generality we may assume that rectifiable curves are Lipschitz continuous.

A *length space* is a metric space such that the distance between any two points equals the infimum of lengths of curves connecting these two points and the space is a *geodesic space* if for any two points, there is a curve that connects these points and whose length equals the distance between the two points. Clearly, any geodesic space is a length space. A shortest curve connecting given two points (if it exists) is called a *geodesic*.

A metric space is *proper* if bounded and closed sets are compact. Proper spaces are also known as *boundedly compact* spaces. The following fact is well know [23, Theorem 3.9].

Lemma 26. *If a metric space X is proper, and if given two points $x, y \in X$ can be connected by a rectifiable curve, then there is a shortest curve connecting x to y.*

Corollary 27. *Any proper length space is geodesic. In particular compact length spaces are geodesic.*

The *speed* of a Lipschitz curve $\gamma : [a, b] \to X$ is defined as

$$|\dot{\gamma}|(t) = \lim_{h \to 0} \frac{d(\gamma(t+h), \gamma(t))}{|h|}.$$

The next result is well known, see e.g., [23, Theorem 3.6].

Lemma 28. *If $\gamma : [a, b] \to X$ is Lipschitz, then the speed $|\dot{\gamma}|(t)$ exists for almost all $t \in [a, b]$ and*

$$\ell(\gamma) = \int_a^b |\dot{\gamma}|(t)\, dt.$$

2.5 Metric trees

The next result provides several equivalent conditions. A metric space that satisfies any of these conditions is called a *metric tree* or an $\mathbb{R}$-*tree*.

Lemma 29. *Let X be a geodesic space. Then the following conditions are equivalent.*

(a) *For any $x, y \in X, x \neq y$, there is a unique arc with endpoints x and y.*

(b) *No subset of X is homeomorphic to $\mathbb{S}^1$.*

(c) *X is simply connected and $\dim X = 1$ (see Section 6.4).*

(d) *Every geodesic triangle is isometric to a tripod.*

(e) *X is 0-hyperbolic in the sense of Gromov.*

(f) *Intersection of any two closed balls is a closed ball or an empty set.*

(g) *For all Lipschitz maps $\gamma : \mathbb{S}^1 \to X$ and $\pi : X \to \mathbb{R}^2$, we have*

$$\int_{\mathbb{S}^1} (\pi \circ \gamma)^*(x\, dy) = 0.$$

A subset of a metric space is called an *arc* if it is homeomorphic to the interval [0,1]. The *endpoints* of an arc, are the images of 0 and 1. A *tripod* is a geodesic space consisting of three segments that meet at one point. We will not recall the definition of the Gromov hyperbolic space since we will not use it in this thesis. We collected the equivalent conditions for reader's convenience and in fact we will mainly need condition (g).

For equivalence between (d), (e) and (f) and (g), see [48]. For equivalence of (a) and (b) see [6, Proposition 2.3]. Finally, the equivalence between (a) and (c) and (e) can be found in [3].

2.6 The Heisenberg groups

Material of this section will be only used in the proof of Theorem 126, which is an application of Theorem 111, and will not play any role in the other parts of the thesis, so the reader may skip this section.

For any positive integer n, we define the Heisenberg group as $\mathbb{H}^n = \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{2n+1}$, with the group law defined by

$$(x, y, t) * (x', y', t') = (x + x', t + y', t + t' + 2\sum_{j=1}^{n}(y_j x'_j - x_j y'_j)).$$

This is a Lie group and as a basis of left invariant vector fields we can take at any given point $(x_1, y_1, \dots, x_n, y_n, t) \in \mathbb{H}^n$ the vectors

$$X_j = \frac{\partial}{\partial x_j} + 2y_j \frac{\partial}{\partial t}, \quad Y_j = \frac{\partial}{\partial y_j} - 2x_j \frac{\partial}{\partial t}, \quad T = \frac{\partial}{\partial t} \quad j = 1, 2, \dots, n.$$

The Heisenberg group is equipped with the left invariant Riemannian metric $\mathfrak{g}$ such that the the vector fields X_j, Y_j, T are orthonormal.

The Heisenberg group is equipped with the so called *horizontal distribution*

$$H_p\mathbb{H}^n = \operatorname{span}\{X_1(p), Y_1(p), \dots, X_n(p), Y_n(p)\} \quad \text{for all } p \in \mathbb{H}^n.$$

This is a smooth distribution of $2n$-dimensional subspaces in the $(2n+1)$-dimensional tangent space $T_p\mathbb{H}^n = T_p\mathbb{R}^{2n+1}$. A vector $v \in T_p\mathbb{R}^{2n+1}$ is *horizontal* if $v \in H_p\mathbb{H}^n$.

An Lipschitz curve $\gamma : [a, b] \to \mathbb{R}^{2n+1}$ is a *horizontal curve* if it is almost everywhere tangent to the horizontal distribution i.e., $\gamma'(t) \in H_{\gamma(t)}\mathbb{H}^n$ for almost every $t \in [a, b]$. It is well known that any two points in $\mathbb{H}^n$ can be connected by a horizontal curve. The *Carnot-Carathéodory metric* d_{cc} in $\mathbb{H}^n$ is defined as the infimum of lengths (computed with respect to the metric $\mathfrak{g}$) of horizontal curves connecting given two points. When we talk about the Heisenberg group, we always regard it as a metric space with the Carnot-Carathéodory metric d_{cc}. The length of a rectifiable curve γ in $(\mathbb{H}^n, d_{cc})$ will be denoted by $\ell_{cc}(\gamma)$.

Let $\pi : \mathbb{R}^{2n+1} \to \mathbb{R}^{2n}$, $\pi(x, y, t) = (x, y)$ be the projection onto the first $2n$-coordinates. The next result is well known.

Lemma 30. *If $\gamma : [a, b] \to \mathbb{H}^n$ is Lipschitz continuous, then $\pi \circ \gamma : [a, b] \to \mathbb{R}^{2n}$ is Lipschitz continuous and $\ell_{cc}(\gamma) = \ell(\pi \circ \gamma)$.*

In other words, the projection $\pi : \mathbb{H}^n \to \mathbb{R}^{2n}$ preserves lengths of Lipschitz curves.

We will need the following nontrivial Lipschitz extension result in the proof of Theorem 126. It is a corollary to [49, Theorem 1.2].

Lemma 31 (Wenger-Young)**.** *If $k \leq n$, then the the pair $(\mathbb{R}^k, \mathbb{H}^n)$ has the Lipschitz extension property, i.e. there is a constant $C > 0$ such that for any $A \subset \mathbb{R}^k$ and any L-Lipschitz map $f\colon A \to \mathbb{H}^n$, there is a CL-Lipschitz map $F\colon \mathbb{R}^k \to \mathbb{H}^n$ satisfying $F(x) = f(x)$ for all $x \in A$.*

2.7 Miscellaneous

The following lemma follows from the universal fact that $\operatorname{diam}(f(A_i)) \leq (\operatorname{Lip} f) \operatorname{diam} A_i$.

Lemma 32. *Let $f\colon X \to Y$ be an L-Lipschitz map, $L > 0$, between metric spaces, then for any $A \subset X$, and any $\delta \in (0, \infty]$,*

$$\mathcal{H}^\alpha_\delta(f(A)) \leq (L)^\alpha \ \mathcal{H}^\alpha_{L\delta}(A)\,.$$

Remark 33. The more interesting cases is when $\delta = \infty$ and the limit as $\delta \to 0$, which yields

$$\mathcal{H}^\alpha(f(A)) \leq (L)^\alpha \ \mathcal{H}^\alpha(A)\,.$$

The following is the famous Rademacher's differentiability theorem.

Theorem 34. *Any Lipschitz map $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is differentiable almost everywhere.*

Lemma 35 (Lusin)**.** *Let X be a metric space and μ and Borel measure on X such that X is a union of finitely many open sets of finite measure. If $\Phi : X \to Y$ is a Borel measurable mapping into a separable metric space Y, then for any $\varepsilon > 0$, there is a closed set $F \subset X$ such that $\mu(X \setminus F) < \varepsilon$ and $\Phi|_F : F \to Y$ is continuous.*

Let $A \subset \mathbb{R}^n$ be measurable. We say that $x \in \mathbb{R}^n$ is a *a density point* of A if

$$\lim_{r \to 0^+} \frac{\mathcal{L}^n(A \cap B(x,r))}{\omega_n r^n} = 1\,.$$

The following corollary of the Lebesgue differentiation theorem is well-known.

Theorem 36. *For a measurable $A \subset \mathbb{R}^n$, almost every point of A is its density point.*

Throughout this work, ℓ^∞ will denote the space of all bounded real sequences $(a_1, a_2, \cdots)$. It is a Banach space when equipped with the norm

$$\|(a_1, a_2, \cdots)\|_{\ell^\infty} = \sup_i |a_i|.$$

Metric spaces do not come with any linear structure on them. However, we have the following important result that we will use extensively. Recall that a metric space is called *separable* if it contains a countable dense subset.

Lemma 37 (Kuratowski-Fréchet)**.** *Every separable metric space admits an isometric embedding into ℓ^∞.*

Indeed, if $x_o \in X$ and $\{x_i\}_{i=1}^\infty \subset X$ is a dense subset, then

$$X \ni x \mapsto \kappa(x) = \big(d(x, x_i) - d(x_i, x_o)\big)_{i=1}^\infty \in \ell^\infty$$

is an isometric embedding.

We denote by ℓ^∞_m the space $\mathbb{R}^m$ with the norm $\|v\|_\infty = \max_{1 \le i \le m} |v_i|$, $v = (v_1, \ldots, v_m)$. There is an obvious projection $\ell^\infty \to \ell^\infty_m$.

3.0 Weighted integrals and weighted Hausdorff measure

In practice it is not easy to compute upper integrals using their definition. In this chapter we give an equivalent characterization of upper integrals for the special case of upper integrals with respect to Hausdorff measure on metric spaces. The new characterization is very flexible and plays a key role in the proof of the coarea inequalities in Chapter 2. However, being objects of independent interest in their own right, we present them separately here.

Please refer to Section 1.2 for historical comments.

Throughout this chapter (X, d) will be a metric space and functions $f : X \to [0, \infty]$ will not necessarily be measurable.

Definition 38. For a function $f : X \to [0, \infty]$, a *weighted covering of* f is a countable collection $\{(a_i, A_i)\}_{i\in\mathbb{N}}$ of pairs of bounded sets $A_i \subset X$ and numbers $a_i \in [0, \infty]$ such that

$$f(x) \leq \sum_i a_i \chi_{A_i}(x) \quad \text{for all} \quad x \in X\,. \tag{3.1}$$

If in addition $\operatorname{diam} A_i \leq \delta$, $\delta \in (0, +\infty]$, for all $i \in \mathbb{N}$, we say that $\{(a_i, A_i)\}_{i\in\mathbb{N}}$ is a *weighted δ-covering of f*. If $f = \chi_E$ we call $\{(a_i, A_i)\}_{i\in\mathbb{N}}$ a *weighted (δ-)covering of E.*

Let $\delta \in (0, +\infty]$, and $s \in [0, \infty)$. The *weighted integral of f* is defined by

$$\int_X^{\bullet} f \, d\mathcal{H}^s_\delta := \inf \sum_{i=1}^{\infty} a_i \zeta^s(A_i), \tag{3.2}$$

where the infimum is taken over all weighted δ-coverings of f, and

$$\int_X^{\bullet} f \, d\mathcal{H}^s = \lim_{\delta\to 0^+} \int_X^{\bullet} f \, d\mathcal{H}^s_\delta.$$

Note that the limit exists since the integral (3.2) is non-increasing in δ.

If no δ-cover of f exists, we set the weighted integral of f to be $+\infty$.

Remark 39. Since the diameter of a set and of its closure are equal, we may assume that the sets A_i are closed.

Definition 40. The *weighted Hausdorff content* and the *weighted Hausdorff measure* of a set $E \subset X$ are respectively defined by

$$\lambda^s_\delta(E) = \int_X^\bullet \chi_E \, d\mathcal{H}^s_\delta \quad \text{and} \quad \lambda^s(E) = \lim_{\delta \to 0^+} \lambda^s_\delta(E) = \int_X^\bullet \chi_E \, d\mathcal{H}^s.$$

In other words $\lambda^s_\delta(E) = \inf \sum_{i=1}^\infty a_i \zeta^s(A_i)$, where the infimum is taken over all collections $\{(a_i, A_i)\}_{i\in\mathbb{N}}$ such that $\sum a_i \chi_{A_i}(x) \geq 1$ for all $x \in E$, and $\operatorname{diam} A_i \leq \delta$, for all $i \in \mathbb{N}$.

Remark 41. Note that while in the definition of a step function we assumed that the sets A_i were disjoint, the sets A_i here are not required to be disjoint. A step function uniquely determines the sets A_i and numbers a_i, but the same function on the right hand side of (3.1) can be represented in several different ways. It is important that the infimum in (3.2) is taken over all collections $\{(a_i, A_i)\}$ and not only over those corresponding to step functions.

Remark 42. It seems that Federer [22, 2.10.24] was the first to define weighted integrals. He denoted them by $\lambda_\delta(f)$ but did not use any terms to refer to them. The first systematic study of weighted measures was done by Kelly [33, 32] under the name of *method III measures*, although he is using the name weighted covering. The name weighted Hausdorff measures was introduced by Howroyd [29], see also [41, Chapter 8]. The term weighted integral and the notation $\int_X^\bullet f \, d\mathcal{H}^s_\delta$ appears in [45].

3.1 Fundamental properties of weighted integrals

Theorem 43. *Let X be a metric space and $s \in [0, \infty)$. Then for any $E \subset X$,*

$$\lambda^s(E) = \mathcal{H}^s(E). \tag{3.3}$$

Moreover, if $\delta \in (0, \infty]$, then

$$(8 \cdot 6^s)^{-1} \mathcal{H}^s_{6\delta}(E) \leq \lambda^s_\delta(E) \leq \mathcal{H}^s_\delta(E). \tag{3.4}$$

Remark 44. Passing to the limit in (3.4) as $\delta \to 0^+$, yields $(8 \cdot 6^s)^{-1} \mathcal{H}^s(E) \leq \lambda^s(E) \leq \mathcal{H}^s(E)$ which is weaker than (3.3) so (3.3) is somewhat surprising.

Theorem 43 will play a crucial role in the proof of

Theorem 45. *Let X be a metric space. For $s \in [0, \infty)$, and any $f : X \to [0, \infty]$ we have*

$$\int_X^{\bullet} f \, d\mathcal{H}^s = \int_X^{*} f \, d\mathcal{H}^s \, . \tag{3.5}$$

Remark 46. Inequality (3.4) is stated implicitly in [22, 2.10.24], as a step in the proof of Theorem 45 (under assumptions (a') or (b')) and the general case follows from the theorem of Davies [10], see [29, 33, 32].

3.2 Weighted covering theorem

The proof of inequality (3.4) is based on the following weighted covering result that we learned from Nazarov through MathOverflow [38]. The result is interesting on its own and we believe it will have applications beyond those given here.

Theorem 47. *Let E be a bounded and non-empty subset of a metric space. If $0 \le b_i < \infty$, $i = 1, 2, \ldots, N$, are fixed numbers and $\{(a_i, B_i)\}_{i=1}^N$ is a finite weighted covering of E by (either all open or all closed) balls i.e.,*

$$\chi_E \le \sum_{i=1}^N a_i \chi_{B_i}, \quad a_i \ge 0, \tag{3.6}$$

then there is a subfamily of pairwise disjoint balls $\{B_{i_j}\}_{j=1}^k$ such that

$$E \subset \bigcup_{j=1}^k 3B_{i_j} \quad \textit{and} \quad \sum_{j=1}^k b_{i_j} \le 2 \sum_{i=1}^N a_i b_i .$$

Remark 48. Later, we will apply Theorem 47 with $b_i = \zeta^s(B_i)$.

Proof. We will prove the result using induction with respect to N. More precisely, we will prove that for every $N \in \mathbb{N}$, the statement is true for any set E that is bounded and non-empty and any weighted covering of it with N balls.

It is important to prove the statement for all sets E. Proving it for a fixed set E would not work, since the induction hypothesis will be applied to sets different than E. Namely, it will be applied to subsets of E.

If $N = 1$, the claim is obvious, because we have one ball B_1 and $a_1 \geq 1$. Suppose $N \geq 2$ and the claim is true if the number of balls is less than or equal to $N - 1$, we will prove it for N balls.

Let $\{(a_i, B_i)\}_{i=1}^N$ be a weighted covering of E satisfying (3.6). For $\alpha = (\alpha_1, \dots, \alpha_N)$, let

$$W = \Big\{\alpha : \alpha_i \geq 0, \ \sum_{i=1}^N \alpha_i \chi_{B_i} \geq \chi_E\Big\}, \quad W_c = \Big\{\alpha : 1 \geq \alpha_i \geq 0, \ \sum_{i=1}^N \alpha_i \chi_{B_i} \geq \chi_E\Big\}.$$

Let $\psi(\alpha) = \sum_{i=1}^N \alpha_i b_i$. If $\alpha \in W$, then

$$\alpha \wedge 1 = (\min\{\alpha_1, 1\}, \dots, \min\{\alpha_N, 1\}) \in W_c \quad \text{and} \quad \psi(\alpha \wedge 1) \leq \psi(\alpha)$$

so $\inf_W \psi = \inf_{W_c} \psi$. Since W_c is compact and non-empty, there is $\alpha \in W_c$ such that $\psi(\alpha) = \inf_{W_c} \psi = \inf_W \psi$. In particular,

$$\sum_{i=1}^N \alpha_i b_i \leq \sum_{i=1}^N a_i b_i. \tag{3.7}$$

If there is i_o such that $\alpha_{i_o} = 0$, we are done. Indeed,

$$\chi_E \leq \sum_{\substack{1 \leq i \leq N \\ i \neq i_o}} \alpha_i \chi_{B_i}$$

is a weighted covering of E by $N - 1$ balls so according to the induction hypothesis, there is a subfamily of pairwise disjoint balls $\{B_{i_j}\}_{j=1}^k$ such that

$$E \subset \bigcup_{j=1}^k 3B_{i_j}, \quad \sum_{j=1}^k b_{i_j} \leq 2 \sum_{\substack{1 \leq i \leq N \\ i \neq i_o}} \alpha_i b_i = 2 \sum_{i=1}^N \alpha_i b_i \leq 2 \sum_{i=1}^N a_i b_i\,.$$

Therefore, we may assume that $\alpha_i > 0$ for all $i \in \{1, \dots, N\}$.

Lemma 49. *If $\alpha \in W_c$ is a minimizer of ψ and $\alpha_i > 0$ for all i, then for any $i_1 \in \{1, \dots, N\}$, we have*

$$\sum_{\{i \colon B_i \cap B_{i_1} \neq \varnothing\}} \alpha_i b_i \geq \frac{b_{i_1}}{2} .$$

Proof. Since the sum on the left hand side includes $\alpha_{i_1} b_{i_1}$, the claim is obvious if $\alpha_{i_1} \geq 1/2$. Therefore, we may assume that $0 < \alpha_{i_1} < 1/2$. Let $0 < h < \alpha_{i_1}$ and define

$$\tilde{\alpha}_i = \begin{cases} \alpha_i & \text{if } B_i \cap B_{i_1} = \varnothing, \\ \alpha_i(1+2h) & \text{if } B_i \cap B_{i_1} \neq \varnothing,\ i \neq i_1, \\ \alpha_i - h & \text{if } i = i_1. \end{cases}$$

We claim that

$$(\tilde{\alpha}_1, \dots, \tilde{\alpha}_N) \in W \quad \text{i.e.,} \quad \sum_{i=1}^{N} \tilde{\alpha}_i \chi_{B_i} \geq \chi_E. \tag{3.8}$$

If $x \notin B_{i_1}$, then $\tilde{\alpha}_{i_1} \chi_{B_{i_1}}(x) = \alpha_{i_1} \chi_{B_{i_1}}(x) = 0$. Since $\tilde{\alpha}_i \geq \alpha_i$ for all $i \neq i_1$, we have

$$\sum_{i=1}^{N} \tilde{\alpha}_i \chi_{B_i}(x) \geq \sum_{i=1}^{N} \alpha_i \chi_{B_i}(x) \geq \chi_E(x). \tag{3.9}$$

If $x \notin E$, then $\chi_E(x) = 0$ and there is nothing to prove.

If $x \in E \cap B_{i_1}$, then

$$1 = \chi_E(x) \leq \sum_{i=1}^{N} \alpha_i \chi_{B_i}(x) = \alpha_{i_1} + \sum_{\{i \colon i \neq i_1,\ x \in B_i \cap B_{i_1}\}} \alpha_i,$$

and hence

$$\sum_{\{i \colon i \neq i_1,\ x \in B_i \cap B_{i_1}\}} \alpha_i \geq 1 - \alpha_{i_1}.$$

Therefore,

$$\begin{aligned} \sum_{i=1}^{N} \tilde{\alpha}_i \chi_{B_i}(x) &= (\alpha_{i_1} - h) + \sum_{\{i \colon i \neq i_1,\ x \in B_i \cap B_{i_1}\}} \alpha_i (1+2h) \\ &\geq (\alpha_{i_1} - h) + (1+2h)(1 - \alpha_{i_1}) = 1 + h(1 - 2\alpha_{i_1}) > 1 = \chi_E(x), \end{aligned}$$

where the last inequality is a consequence of $0 < \alpha_{i_1} < 1/2$. This completes the proof of (3.8).

Since ψ attains minimum at α, we have

$$\sum_{i=1}^{N} \alpha_i b_i \le \sum_{i=1}^{N} \tilde{\alpha}_i b_i. \tag{3.10}$$

Since $\alpha_i = \tilde{\alpha}_i$ if $B_i \cap B_{i_1} = \varnothing$, (3.10) yields

$$\begin{aligned} &\alpha_{i_1} b_{i_1} + \sum_{\{i:\, i\neq i_1,\ B_i\cap B_{i_1}\neq\varnothing\}} \alpha_i b_i \\ &\le (\alpha_{i_1} - h) b_{i_1} + \sum_{\{i:\, i\neq i_1,\ B_i\cap B_{i_1}\neq\varnothing\}} \alpha_i (1+2h) b_i, \end{aligned}$$

and hence

$$h b_{i_1} \le 2h \sum_{\{i:\, i\neq i_1,\ B_i\cap B_{i_1}\neq\varnothing\}} \alpha_i b_i$$

which finishes the proof of Lemma 49. □

Now we can complete the proof of the theorem. Let B_{i_1} be a ball with the largest diameter and let

$$I = \{i :\ B_i \cap B_{i_1} \neq \varnothing\} \quad \text{and} \quad I^c = \{i :\ B_i \cap B_{i_1} = \varnothing\}.$$

We have

$$\bigcup_{i\in I} B_i \subset 3B_{i_1} \quad \text{and} \quad \sum_{i\in I} \alpha_i b_i \ge \frac{b_{i_1}}{2}.$$

The inclusion is a consequence of the triangle inequality and the fact that $\operatorname{diam} B_{i_1} \ge \operatorname{diam} B_i$ for $i \in I$, while the inequality follows from Lemma 49.

If $E \setminus 3B_{i_1} = \varnothing$, then (3.7) yields

$$E \subset 3B_{i_1} \quad \text{and} \quad b_{i_1} \le 2\sum_{i\in I} \alpha_i b_i \le 2\sum_{i=1}^{N} \alpha_i b_i \le 2\sum_{i=1}^{N} a_i b_i\,,$$

and the theorem follows.

Therefore, we may assume that $E \setminus 3B_{i_1} \neq \varnothing$. Since the balls B_i, $i \in I$ have empty intersection with $E \setminus 3B_{i_1}$,

$$\sum_{i\in I^c} \alpha_i \chi_{B_i} \ge \chi_{E\setminus 3B_{i_1}}$$

and hence $\{(\alpha_i, B_i)\}_{i\in I^c}$ is a weighted covering of $E \setminus 3B_{i_1}$ and the number of balls in that covering is less than or equal to $N-1$ (we removed at least one ball: B_{i_1}). According to the induction hypothesis, we can select pairwise disjoint balls $\{B_{i_j}\}_{j=2}^k$, $i_j \in I^c$ such that

$$E \setminus 3B_{i_1} \subset \bigcup_{j=2}^k 3B_{i_j} \quad \text{and} \quad \sum_{j=2}^k b_{i_j} \le 2 \sum_{i\in I^c} \alpha_i b_i\,.$$

Therefore,

$$E \subset 3B_{i_1} \cup \bigcup_{j=2}^k 3B_{i_j} = \bigcup_{j=1}^k 3B_{i_j}$$

(note that $B_{i_1} \cap B_{i_j} = \varnothing$, for $j \ge 2$ so the balls $\{B_{i_j}\}_{j=1}^k$ are pairwise disjoint) and

$$\sum_{j=1}^k b_{i_j} = b_{i_1} + \sum_{j=2}^k b_{i_j} \le 2\sum_{i\in I} \alpha_i b_i + 2\sum_{i\in I^c} \alpha_i b_i = 2\sum_{i=1}^N \alpha_i b_i \le \sum_{i=1}^N a_i b_i\,.$$

The proof is complete. □

Corollary 50. *Let E be a non-empty subset of a metric space, $\{b_i\}_{i=1}^\infty$, a sequence of non-negative numbers, and $\{(a_i, B_i)\}_{i=1}^\infty$, a weighted covering of E by (all open or all closed) balls i.e.,*

$$\chi_E \le \sum_{i=1}^\infty a_i \chi_{B_i}, \quad a_i \ge 0.$$

Then there is a subfamily of balls $\{B_{i_j}\}_{j=1}^\infty$ such that

$$E \subset \bigcup_{j=1}^\infty 3B_{i_j} \quad \textit{and} \quad \sum_{j=1}^\infty b_{i_j} \le 8 \sum_{i=1}^\infty a_i b_i.$$

Remark 51. Differently than in Theorem 47, we do not assume that the balls $\{B_{i_j}\}_{j=1}^\infty$ are pairwise disjoint.

Proof. If $\sum_{i=1}^{\infty} a_i b_i = +\infty$, the claim is obvious. Therefore, we may assume that $M := \sum_{i=1}^{\infty} a_i b_i < \infty$. We divide the series into finite blocks such that

$$\sum_{i=1}^{\infty} a_i b_i = \sum_{k=0}^{\infty} \Big(\underbrace{\sum_{i=N_k+1}^{N_{k+1}} a_i b_i}_{\leq 4^{-k}M} \Big), \quad 0 = N_0 < N_1 < N_2 < \ldots$$

Let

$$E_k = \left\{ x \in E : \sum_{i=N_k+1}^{N_{k+1}} 2^{k+1} a_i \chi_{B_i}(x) \geq 1 \right\}.$$

Observe that $E = \bigcup_{k=0}^{\infty} E_k$. Indeed, if $x \in E$, then

$$\sum_{k=0}^{\infty} \Big(\sum_{i=N_k+1}^{N_{k+1}} a_i \chi_{B_i}(x) \Big) = \sum_{i=1}^{\infty} a_i \chi_{B_i}(x) \geq \chi_E(x) = 1 = \sum_{k=0}^{\infty} 2^{-(k+1)}.$$

Therefore, there is k such that

$$\sum_{i=N_k+1}^{N_{k+1}} a_i \chi_{B_i}(x) \geq 2^{-(k+1)}, \quad \text{so} \quad x \in E_k.$$

By the definition of E_k, the family $\{(2^{k+1}a_i, B_i)\}_{i=N_k+1}^{N_{k+1}}$ is a finite weighted covering of E_k. According to Theorem 47, we can select pairwise disjoint balls $\{B_{i_j}^{(k)}\}_{j=1}^{\ell_k}$ from $\{B_i\}_{i=N_k+1}^{N_{k+1}}$ so that

$$E_k \subset \bigcup_{j=1}^{\ell_k} 3B_{i_j}^{(k)} \quad \text{and} \quad \sum_{j=1}^{\ell_k} b_{i_j}^{(k)} \leq 2 \sum_{i=N_k+1}^{N_{k+1}} 2^{k+1} a_i b_i < 4 \cdot 2^{-k} M.$$

To be more precise, we select this family of balls only if $E_k \neq \varnothing$. If $E_k = \varnothing$, we select empty family of balls.

If we relabel balls as

$$\{B_{i_j}^{(k)} : \, k \in \mathbb{N} \cup \{0\}, \; 1 \leq j \leq \ell_k\} := \{B_{i_j}\}_{j=1}^{\infty},$$

then

$$E = \bigcup_{k=0}^{\infty} E_k \subset \bigcup_{k=0}^{\infty} \bigcup_{j=1}^{\ell_k} 3B_{i_j}^{(k)} = \bigcup_{j=1}^{\infty} 3B_{i_j},$$

and

$$\sum_{j=1}^{\infty} b_{i_j} = \sum_{k=0}^{\infty} \sum_{j=1}^{\ell_k} b_{i_j}^{(k)} \leq \sum_{k=0}^{\infty} 4 \cdot 2^{-k} M = 8M = 8 \sum_{i=1}^{\infty} a_i b_i.$$

$\square$

3.3 Proof of Theorem 43

First we will prove (3.4). Note that the inequality $\lambda^s_\delta(E) \le \mathcal{H}^s_\delta(E)$ is obvious and follows upon taking weighted coverings with coefficients $a_i = 1$ so it remains to prove that $\mathcal{H}^s_{6\delta}(E) \le 8 \cdot 6^s \lambda^s_\delta(E)$.

Let $\{(a_i, A_i)\}_{i=1}^\infty$ be a weighted δ-covering of E,

$$\chi_E \le \sum_{i=1}^\infty a_i \chi_{A_i}, \quad a_i \ge 0, \quad \operatorname{diam} A_i \le \delta.$$

Each of the sets A_i is contained in a closed ball B_i of radius $\operatorname{diam} A_i$. Hence

$$\operatorname{diam}(3B_i) \le 6 \operatorname{diam} A_i \le 6\delta \quad \text{so} \quad \zeta^s(3B_i) \le 6^s \zeta^s(A_i).$$

Since $\{(a_i, B_i)\}_{i=1}^\infty$ is also a weighted cover of E, Corollary 50 with $b_i = \zeta^s(A_i)$ yields a subfamily $\{B_{i_j}\}_{j=1}^\infty$ of balls such that

$$E \subset \bigcup_{j=1}^\infty 3B_{i_j} \quad \text{and} \quad \sum_{j=1}^\infty \zeta^s(A_{i_j}) \le 8 \sum_{i=1}^\infty a_i \zeta^s(A_i).$$

Therefore,

$$\mathcal{H}^s_{6\delta}(E) \le \sum_{j=1}^\infty \zeta^s(3B_{i_j}) \le 6^s \sum_{j=1}^\infty \zeta^s(A_{i_j}) \le 8 \cdot 6^s \sum_{i=1}^\infty a_i \zeta^s(A_i)$$

and taking the infimum over all weighted δ-coverings $\{(a_i, A_i)\}_{i=1}^\infty$ of E proves that $\mathcal{H}^s_{6\delta}(E) \le 8 \cdot 6^s \lambda^s_\delta(E)$ and completes the proof of (3.4).

Passing to the limit in (3.4) as $\delta \to 0^+$ yields

$$(8 \cdot 6^s)^{-1} \mathcal{H}^s(E) \le \lambda^s(E) \le \mathcal{H}^s(E).$$

This proves (3.3) when $\mathcal{H}^s(E) = \infty$. Therefore, it remains to prove

$$\mathcal{H}^s(E) \le \lambda^s(E) \quad \text{assuming that } \mathcal{H}^s(E) < \infty. \tag{3.11}$$

Let $\tilde{E}$ be a Borel set such that $E \subset \tilde{E}$ and $\mathcal{H}^s(\tilde{E}) = \mathcal{H}^s(E)$.

Fix $\varepsilon > 0$. For each $j \in \mathbb{N}$, let W_j be the set of points $x \in \tilde{E}$ such that

$$x \in F \subset \bar{B}(x, 1/j) \quad \Longrightarrow \quad \mathcal{H}^s(\tilde{E} \cap F) \le (1+\varepsilon)\zeta^s(F).$$

Note that $W_1 \subset W_2 \subset \dots$ and Lemma 24 applied to $\tilde{E}$ regarded as a metric space yields (because $\mathcal{H}^s(\tilde{E}) < \infty$)

$$\mathcal{H}^s\Big(\tilde{E} \setminus \bigcup_{j=1}^{\infty} W_j\Big) = 0.$$

Therefore, Lemma 23 implies

$$\mathcal{H}^s(E) \leq \mathcal{H}^s\Big(E \cap \bigcup_{j=1}^{\infty} W_j\Big) + \underbrace{\mathcal{H}^s\Big(E \setminus \bigcup_{j=1}^{\infty} W_j\Big)}_{0} = \lim_{j\to\infty} \mathcal{H}^s(E \cap W_j).$$

It remains to show that

$$\mathcal{H}^s(E \cap W_j) \leq (1+\varepsilon)(\lambda^s_{1/j}(E) + \varepsilon) \tag{3.12}$$

as passing to the limit as $j \to \infty$ and then as $\varepsilon \to 0^+$ will imply (3.11).

Fix $j \in \mathbb{N}$. Let $\{(a_k, A_k)\}_{k=1}^{\infty}$ be a weighted $1/j$-covering of E by closed sets such that

$$\sum_{k=1}^{\infty} a_k \zeta^s(A_k) \leq \lambda^s_{1/j}(E) + \varepsilon. \tag{3.13}$$

Let $I = \{k : W_j \cap A_k \neq \varnothing\}$. We have

$$\chi_{E\cap W_j} \leq \sum_{k\in I} a_k \chi_{\tilde{E}\cap A_k}.$$

Let

$$Z = \Big\{x : \sum_{k\in I} a_k \chi_{\tilde{E}\cap A_k}(x) \geq 1\Big\}.$$

The set Z is Borel, $E \cap W_j \subset Z$, and

$$\chi_Z \leq \sum_{k\in I} a_k \chi_{\tilde{E}\cap A_k}.$$

Integrating this inequality with respect to $\mathcal{H}^s$ yields

$$\mathcal{H}^s(E \cap W_j) \leq \mathcal{H}^s(Z) \leq \sum_{k\in I} a_k \mathcal{H}^s(\tilde{E} \cap A_k).$$

If $k \in I$, then there is $x \in W_j \cap A_k$ and hence

$$x \in A_k \subset \bar{B}(x, 1/j) \quad \text{so} \quad \mathcal{H}^s(\tilde{E} \cap A_k) \leq (1+\varepsilon)\zeta^s(A_k)$$

by the definition of the set W_j. Therefore,

$$\mathcal{H}^s(E \cap W_j) \le (1+\varepsilon) \sum_{k \in I} a_k \zeta^s(A_k) \le (1+\varepsilon)(\lambda^s_{1/j}(E) + \varepsilon),$$

where the last inequality follows from (3.13). This proves (3.12) and completes the proof of the theorem.

□

3.4 Proof of Theorem 45

We first prove the following easier inequality

$$\int_X^{\bullet} f \; d\mathcal{H}^s \le \int_X^* f \; d\mathcal{H}^s \,. \tag{3.14}$$

To this end it suffices to prove that for any $\delta > 0$

$$\int_X^{\bullet} f \; d\mathcal{H}^s_\delta \le \int_X^* f \; d\mathcal{H}^s \,, \tag{3.15}$$

as (3.14) will follow upon passing to the limit as $\delta \to 0^+$. Assume that the right-hand side of (3.15) is finite. Given $\varepsilon > 0$, it follows from Lemma 13 that there is a step function

$$f \le \sum_{i=1}^\infty a_i \chi_{A_i}, \quad a_i > 0$$

such that

$$\sum_{i=1}^\infty a_i \mathcal{H}^s_\delta(A_i) \le \sum_{i=1}^\infty a_i \mathcal{H}^s(A_i) \le \int_X^* f \, d\mathcal{H}^s + \frac{\varepsilon}{2}.$$

For each i, there is a δ-covering $A_i \subset \bigcup_{j=1}^\infty A_{ij}$, satisfying

$$\sum_{j=1}^\infty \zeta^s(A_{ij}) < \mathcal{H}^s_\delta(A_i) + \frac{\varepsilon}{2^{i+1} a_i}.$$

Then with $a_{ij} = a_i$,

$$f \le \sum_{i=1}^\infty a_i \chi_{A_i} \le \sum_{i=1}^\infty a_i \Big(\sum_{j=1}^\infty \chi_{A_{ij}} \Big) = \sum_{i,j=1}^\infty a_{ij} \chi_{A_{ij}}$$

so $\{(a_{ij},A_{ij})\}_{i,j=1}^{\infty}$ is a weighted δ-covering of f and hence

$$\begin{aligned}\int_X^{\bullet} f\,d\mathcal{H}^s_\delta &\le \sum_{i,j=1}^{\infty} a_{ij}\zeta^s(A_{ij}) = \sum_{i=1}^{\infty} a_i\Big(\sum_{j=1}^{\infty}\zeta^s(A_{ij})\Big)\\ &< \sum_{i=1}^{\infty} a_i\Big(\mathcal{H}^s_\delta(A_i)+\frac{\varepsilon}{2^{i+1}a_i}\Big) = \sum_{i=1}^{\infty} a_i\mathcal{H}^s_\delta(A_i)+\frac{\varepsilon}{2} \le \int_X^{*} f\,d\mathcal{H}^s+\varepsilon.\end{aligned}$$

Since $\varepsilon>0$ was chosen arbitrarily, (3.15) and hence (3.14) follow.

Now we must prove the reverse inequality

$$\int_X^{*} f\,d\mathcal{H}^s \le \int_X^{\bullet} f\,d\mathcal{H}^s\,. \tag{3.16}$$

Clearly, it is important to consider the set $A=\{x\in X: f(x)>0\}$, where the function f is positive. We will split the proof into three cases. We shall also assume that the right-hand side in (3.16) is finite.

Case 1. $\mathcal{H}^s(A)<\infty$.

This case is similar to the proof of (3.3). Let $\varepsilon>0$ be given. According to Theorem 20, there is a Borel set $\tilde{A}$ such that $A\subset\tilde{A}$ and $\mathcal{H}^s(A)=\mathcal{H}^s(\tilde{A})$. Applying Lemma 24 to $\tilde{A}$ regarded as a metric space, we have that there is a set $E\subset\tilde{A}$, $\mathcal{H}^s(E)=0$, such that

$$\forall\, x\in\tilde{A}\setminus E\quad \exists\,\delta_x>0\ \ \forall\, F\subset X\ \ (x\in F\subset\bar{B}(x,\delta_x)\ \Rightarrow\ \mathcal{H}^s(\tilde{A}\cap F)\le(1+\varepsilon)\zeta^s(F)).$$

Let $W_j\subset\tilde{A}$ be the set of points $x\in\tilde{A}$ such that

$$x\in F\subset\bar{B}(x,1/j)\quad\Longrightarrow\quad \mathcal{H}^s(\tilde{A}\cap F)\le(1+\varepsilon)\zeta^s(F). \tag{3.17}$$

Clearly, $W_1\subset W_2\subset\ldots$ and

$$\tilde{A}=E\cup\bigcup_{j=1}^{\infty}W_j,\quad \mathcal{H}^s(E)=0.$$

It suffices to prove that for each j, we have

$$\int_X^{*} f\chi_{W_j}\,d\mathcal{H}^s \le (1+\varepsilon)\left(\int_X^{\bullet} f\,d\mathcal{H}^s_{1/j}+\varepsilon\right), \tag{3.18}$$

because, (3.16) will follow from Lemma 11 upon passing to the limit, first as $j\to\infty$, and then as $\varepsilon\to 0^+$.

According to the definition of the weighted integral and Remark 39, for each j there is a weighted $1/j$-covering

$$f(x) \leq \sum_{k=1}^{\infty} a_k \chi_{A_{jk}}(x), \quad A_{jk}\text{-closed}, \quad \operatorname{diam} A_{jk} \leq \frac{1}{j}$$

such that

$$\sum_{k=1}^{\infty} a_k \zeta^s(A_{jk}) \leq \int_X^\bullet f \, d\mathcal{H}^s_{1/j} + \varepsilon.$$

Let $I = \{k : W_j \cap A_{jk} \neq \varnothing\}$. We have

$$f\chi_{W_j} \leq \sum_{k\in I} a_k \chi_{\tilde{A}\cap A_{jk}}$$

and measurability of the right hand side yields

$$\int_X^* f\chi_{W_j} \, d\mathcal{H}^s \leq \sum_{k\in I} a_k \mathcal{H}^s(\tilde{A}\cap A_{jk}) \leq \heartsuit.$$

If $k \in I$, and $x \in W_j \cap A_{jk}$, then $x \in A_{jk} \subset \bar{B}(x, 1/j)$ so (3.17) yields

$$\heartsuit \leq (1+\varepsilon) \sum_{k\in I} a_k \zeta^s(A_{jk}) \leq (1+\varepsilon)\Big(\int_X^\bullet f \, d\mathcal{H}^s_{1/j} + \varepsilon \Big).$$

This completes the proof of (3.18).

Case 2. $A = \bigcup_{i=1}^\infty A_i$, *where* $\mathcal{H}^s(A_i) < \infty$.

By replacing A_i with $\bigcup_{1\leq j\leq i} A_j$, we can assume further that $A_1 \subset A_2 \subset \ldots$ Since $\mathcal{H}^s(\{x : (f\chi_{A_i})(x) > 0\}) < \infty$, inequality (3.16) follows from Case 1 applied to $f\chi_{A_i}$ and from Lemma 11:

$$\int_X^* f \, d\mathcal{H}^s \xleftarrow{i\to\infty} \int_X^* f\chi_{A_i} \, d\mathcal{H}^s \leq \int_X^\bullet f\chi_{A_i} d\, \mathcal{H}^s \leq \int_X^\bullet f \, d\mathcal{H}^s.$$

Case 3. *The measure $\mathcal{H}^s$ of the set A is not σ-finite.*

In order to prove inequality (3.16), it suffices to show that

$$\int_X^\bullet f \, d\mathcal{H}^s = \infty\,. \tag{3.19}$$

To prove this, we will use Theorem 43. Since the $\mathcal{H}^s$ measure of the set $\{f > 0\}$ is not σ-finite, there is $t > 0$ such that $\mathcal{H}^s(\{f \geq t\}) = \infty$. Therefore, for every $M > 0$, there is $\delta > 0$ such that

$$\mathcal{H}^s_{6\delta}(\{x \in X : f(x) \geq t\}) > M$$

so Theorem 43 yields ($C = 8 \cdot 6^s$):

$$\int_X^\bullet f \, d\mathcal{H}^s \geq \int_X^\bullet t\chi_{\{f \geq t\}} \, d\mathcal{H}^s_\delta = t\lambda^s_\delta(\{f \geq t\}) \geq C^{-1} t \mathcal{H}^s_{6\delta}(\{f \geq t\}) \geq C^{-1} t M \, ,$$

and (3.19) follows. The proof of Theorem 45 is complete.

Acknowledgement. We would like to express our deepest gratitude to Fedor Nazarov for his kindness in providing us with an elementary proof of inequality (3.4), through Math-Overflow [38].

4.0 The coarea inequality

In this chapter we will prove the following theorem, known as the *coarea inequality*, or Eilenberg inequality. In fact, we prove a more general coarea inequality that contains this result as a corollary (see Theorem 58).

Theorem 52. *Let X and Y be arbitrary metric spaces, $0 \leq t \leq s < \infty$ (any) real numbers and $E \subset X$ any subset. Then, for any Lipschitz map $f : X \to Y$ we have*

$$\int_Y^* \mathcal{H}^{s-t}\left(f^{-1}(y) \cap E\right)\, d\mathcal{H}^t(y) \leq (\operatorname{Lip} f)^t\, \frac{\omega_{s-t}\omega_t}{\omega_s} \mathcal{H}^s(E)\,. \tag{4.1}$$

Moreover if X is boundedly compact i.e., bounded and closed sets in X are compact, E is $\mathcal{H}^s$-measurable, and $\mathcal{H}^s(E) < \infty$, then the function

$$y \mapsto \mathcal{H}^{s-t}\left(f^{-1}(y) \cap E\right) \tag{4.2}$$

is $\mathcal{H}^t$-measurable and therefore, the upper integral can be replaced with the usual integral.

Remark 53. In general, we cannot expect measurability of the function (4.2) as the following simple example shows: Let $V \subset \mathbb{R}$ be a non-measurable set. Let $X = V$, $Y = \mathbb{R}$ and $f : X \to Y$, $f(x) = x$. Then for $s = t = 1$, and $E = X$, the function (4.2) is the characteristic function of V and therefore is not measurable. It was communicated to us by Pertti Mattila [40] that (4.2) is measurable with respect to the sigma-algebra generated by analytic sets if X and Y are Polish spaces and E is analytic. This is a consequence of the work of Dellacherie [12], see Remark 7.8 in [41]. However, we did not verify this statement.

Proving measurability of (4.2) under the given assumptions is not difficult, see proof of Theorem 58, so the main focus will be on the inequality.

For the history of the coarea inequality and its proof please see Section 1.2. Our proof follows [19].

4.1 Generalized coarea inequality

The right-hand side in the classical coarea inequality (4.1) is a coarse quantity that only sees the global properties of the map. This is not sufficient for applications that need to take into account the local behavior of the function. In search for a more refined coarea inequality we came up with the following notion of a "mapping content."

Definition 54. For an arbitrary map $f : X \to Y$ between metric spaces, $s, t \in [0, \infty)$, $\delta \in (0, \infty]$, and any $E \subset X$ we define

$$\Phi_\delta^{s,t}(f, E) := \inf \sum_{i=1}^{\infty} \zeta^s(f(A_i))\zeta^t(A_i),$$

where the infimum is taken over all δ-coverings $\{A_i\}_{i=1}^{\infty}$ of E. Obviously, $\delta \mapsto \Phi_\delta^{s,t}$ is non-increasing, allowing the definition

$$\Phi^{s,t}(f, E) := \lim_{\delta \to 0^+} \Phi_\delta^{s,t}(f, E).$$

Remark 55. This definition is motivated by a similar definition in [24, Appendix A] and also by the definition of the mapping content introduced in [4, 8], see Chapter 6.

The proofs of the next two easy results are left to the reader.

Lemma 56. *For any $\delta \in (0, \infty]$, $s, t \in [0, \infty)$, $E, F \subset X$, and $f : X \to Y$ we have*

$$\Phi_\delta^{s,t}(f, E \cup F) \leq \Phi_\delta^{s,t}(f, E) + \Phi_\delta^{s,t}(f, F) \quad \textit{so} \quad \Phi^{s,t}(f, E \cup F) \leq \Phi^{s,t}(f, E) + \Phi^{s,t}(f, F).$$

Lemma 57. *If $f : X \to Y$ is Lipschitz continuous and $E \subset X$, $s, t \in [0, \infty)$, and $\delta \in (0, \infty]$, then*

$$\Phi_\delta^{s,t}(f, E) \leq (\operatorname{Lip} f)^s \frac{\omega_s \omega_t}{\omega_{s+t}} \mathcal{H}_\delta^{s+t}(E) \quad \textit{so} \quad \Phi^{s,t}(f, E) \leq (\operatorname{Lip} f)^s \frac{\omega_s \omega_t}{\omega_{s+t}} \mathcal{H}^{s+t}(E).$$

The next result is a generalization of Theorem 52 and it is motivated by the results in [4, 8, 24].

Theorem 58 (The Generalized Coarea Inequality)**.** *If $f : X \to Y$ is a uniformly continuous map between metric spaces, $0 \le t \le s < \infty$ and $E \subset X$, then*

$$\int_Y^* \mathcal{H}^{s-t}(f^{-1}(y) \cap E)\, d\mathcal{H}^t(y) \le \Phi^{t,s-t}(f, E). \tag{4.3}$$

The strategy to prove the coarea inequality, Theorem 58 (generalization of Theorem 52), is as follows. First we prove a variant of the inequality (4.3) that involves the weighted integral in place of the upper integral (Lemma 59). It turns out that in this case the proof is very simple. Then by Theorem 45 the weighted integral is equal to the upper integral. Finally, we will need to the monotone convergence theorem, Lemma 11. Therefore, it becomes clear that all the difficulties of the proof of the coarea inequality is in proving the equality between weighted integrals and the upper integral, i.e. Theorem 45.

Lemma 59. *If $f : X \to Y$ is a uniformly continuous map between metric spaces, $0 \le t \le s < \infty$ and $E \subset X$, then*

$$\lim_{\delta \to 0^+} \int_Y^\bullet \mathcal{H}_\delta^{s-t}(f^{-1}(y) \cap E)\, d\mathcal{H}^t(y) \le \Phi^{t,s-t}(f, E). \tag{4.4}$$

Remark 60. At this point it is not entirely clear that we can pass to the limit under the sign of the integral as $\delta \to 0^+$, since we do not a priori have the monotone convergence theorem for weighted integrals. In fact such a result is true since according to Theorem 45, the weighted integral equals the upper integral.

4.2 Proofs of the coarea inequalities

Proof of Lemma 59. Assume that $\Phi^{t,s-t}(f, E) < \infty$, as otherwise the inequality is obvious. Fix $\delta_o \in (0, \infty]$. Given $\varepsilon > 0$ and $0 < \delta \le \delta_o$, let $\{A_i\}_{i=1}^\infty$ be a δ-covering of E such that

$$\sum_{i=1}^\infty \zeta^t(f(A_i))\zeta^{s-t}(A_i) < \Phi_\delta^{t,s-t}(f, E) + \varepsilon. \tag{4.5}$$

Since the sets $\{A_i : y \in f(A_i)\}$ form a δ-covering of $f^{-1}(y) \cap E$, we have

$$\mathcal{H}_{\delta_o}^{s-t}(f^{-1}(y) \cap E) \le \sum_{i=1}^\infty a_i \chi_{F_i}(y), \quad \text{where} \quad a_i = \zeta^{s-t}(A_i) \text{ and } F_i = f(A_i). \tag{4.6}$$

Since the mapping f is uniformly continuous,

$$\eta(\delta) = \sup_{\substack{A\subset X \\ \operatorname{diam} A\leq\delta}} \operatorname{diam} f(A) \to 0 \quad \text{as } \delta \to 0^+.$$

According to the inequality (4.6), $\{(a_i, F_i)\}_{i=1}^\infty$ form a weighted $\eta(\delta)$-covering of the function $y \mapsto \mathcal{H}^{s-t}_{\delta_o}(f^{-1}(y) \cap E)$, so, the definition of the weighted integral yields

$$\int_Y^\bullet \mathcal{H}^{s-t}_{\delta_o}(f^{-1}(y) \cap E)\, d\mathcal{H}^t_{\eta(\delta)}(y) \leq \sum_{i=1}^\infty a_i \zeta^t(F_i) < \Phi^{t,s-t}_\delta(f, E) + \varepsilon,$$

where the last inequality is nothing else, but inequality (4.5). Letting $\delta \to 0^+$ first and then $\varepsilon \to 0^+$ proves

$$\int_Y^\bullet \mathcal{H}^{s-t}_{\delta_o}(f^{-1}(y) \cap E)\, d\mathcal{H}^t(y) \leq \Phi^{t,s-t}(f, E).$$

Since, δ_o was arbitrary, (4.4) follows. □

Proof of Theorem 58. It follows immediately from Lemma 59, Theorem 45 and Lemma 11.
□

Proof of Theorem 52. Theorem 58 and Lemma 57 imply inequality (4.1) and it only remains to show measurability of the function (4.2) under the assumptions that X is boundedly compact, E is $\mathcal{H}^s$-measurable and $\mathcal{H}^s(E) < \infty$.

Since bounded and closed sets are compact, Lemma 22 implies existence of a decomposition

$$E = N \cup \bigcup_{i=1}^\infty K_i, \quad \mathcal{H}^s(N) = 0, \quad K_1 \subset K_2 \subset \ldots \text{ compact sets.}$$

It follows from (4.1) that $\mathcal{H}^{s-t}(f^{-1}(y) \cap N) = 0$ for $\mathcal{H}^t$-almost every $y \in Y$ so for almost all $y \in Y$ we have

$$\mathcal{H}^{s-t}(f^{-1}(y) \cap E) = \mathcal{H}^{s-t}\Big(f^{-1}(y) \cap \bigcup_{i=1}^\infty K_i\Big) = \lim_{i\to\infty} \mathcal{H}^{s-t}(f^{-1}(y) \cap K_i).$$

Therefore it remains to show measurability of the function $y \mapsto \mathcal{H}^{s-t}(f^{-1}(y) \cap K)$, where $K \subset X$ is a compact set. To this end it suffices to prove measurability of the sets

$$Y_u = \{y \in Y : \mathcal{H}^{s-t}(f^{-1}(y) \cap K) \leq u\}, \quad u \in \mathbb{R}.$$

If $u < 0$, $Y_u = \varnothing$ so we may assume that $u \geq 0$.

Recall that in Section 2.3 the content $\mathscr{H}^{s-t}_\delta$ was defined with open sets. Since it defines the standard Hausdorff measure, we have

$$Y_u = \bigcap_{j=1}^{\infty} \Big\{ y \in Y : \mathscr{H}^{s-t}_{1/j}(f^{-1}(y) \cap K) < u + \frac{1}{j} \Big\}$$

so it suffices to show that the sets of the form

$$V = \{ y \in Y : \mathscr{H}^{s-t}_\delta(f^{-1}(y) \cap K) < v \}$$

are open (for v and δ positive values). To this end it suffices to show that if $y \in V$ and $y_k \to y$, then $y_k \in V$ for sufficiently large k. For $y \in V$ fix an open covering

$$f^{-1}(y) \cap K \subset \bigcup_{j=1}^{\infty} U_j, \quad \operatorname{diam} U_j < \delta, \quad \sum_{j=1}^{\infty} \zeta^{s-t}(U_j) < v.$$

Using a standard compactness argument, it follows that there exists a k_0 such that $f^{-1}(y_k) \cap K \subset \bigcup_{j=1}^{\infty} U_j$ for $k \geq k_o$ and hence $\mathscr{H}^{s-t}_\delta(f^{-1}(y_k) \cap K) < v$, proving that $y_k \in V$ for $k \geq k_o$. □

4.3 The lower density and doubling spaces

Throughout Section 4.3, X and Y will denote metric spaces. In this section we will improve Theorem 58 under the assumption that the Hausdorff measure on X is doubling. The main result of this section, Theorem 75, is closely related to the coarea formula, see Corollary 77 and Remark 78.

Definition 61. For an arbitrary map $f : E \to Y$, $E \subset X$, $s \in (0, \infty)$, $t \in [0, \infty)$ and $\delta \in (0, \infty]$ we define

$$\tilde{\mathcal{H}}^{s,t}_\delta(f, E) = \inf \sum_{i=1}^{\infty} \mathcal{H}^s_\infty(f(A_i))\zeta^t(A_i),$$

where the infimum is taken over all δ-coverings $\{A_i\}_{i=1}^{\infty}$ of E. If no such covering exists then $\tilde{\mathcal{H}}^{s,t}_\delta(f, E) = \infty$.

The following elementary observation will be useful.

Lemma 62. *For any map $f : E \to Y$, $E \subset X$, $s \in (0,\infty)$, $t \in [0,\infty)$ and $\delta \in (0,\infty]$ we have*

$$\Phi_\delta^{s,t}(f,E) = \tilde{\mathcal{H}}_\delta^{s,t}(f,E).$$

Proof. Since $\mathcal{H}_\infty^s(f(A_i)) \leq \zeta^s(f(A_i))$, the inequality $\tilde{\mathcal{H}}_\delta^{s,t} \leq \Phi_\delta^{s,t}$ is obvious. Therefore, it remains to prove that $\Phi_\delta^{s,t}(f,E) \leq \tilde{\mathcal{H}}_\delta^{s,t}(f,E)$ and we can assume that $\tilde{\mathcal{H}}_\delta^{s,t}(f,E) < \infty$.

Given $\varepsilon > 0$, let $\{A_i\}_{i=1}^\infty$ be a δ-covering of E such that

$$\sum_{i=1}^\infty \zeta^t(A_i)\mathcal{H}_\infty^s(f(A_i)) < \tilde{\mathcal{H}}_\delta^{s,t}(f,E) + \frac{\varepsilon}{2}.$$

For each $i \in \mathbb{N}$, let $\{C_{ij}\}_{j=1}^\infty$ be a covering of $f(A_i)$ such that

$$\sum_{j=1}^\infty \zeta^s(C_{ij}) < \mathcal{H}_\infty^s(f(A_i)) + \frac{\varepsilon}{2^{i+1}(\zeta^t(A_i)+1)}.$$

Let $A_{ij} = A_i \cap f^{-1}(C_{ij})$. Then

$$\begin{aligned}\Phi_\delta^{s,t}(f,E) &\leq \sum_{i,j=1}^\infty \zeta^t(A_{ij})\zeta^s(f(A_{ij})) \leq \sum_{i=1}^\infty \zeta^t(A_i)\Big(\sum_{j=1}^\infty \zeta^s(C_{ij})\Big) \\ &\leq \sum_{i=1}^\infty \zeta^t(A_i)\mathcal{H}_\infty^s(f(A_i)) + \frac{\varepsilon}{2} < \tilde{\mathcal{H}}_\delta^{s,t}(f,E) + \varepsilon\end{aligned}$$

and the result follows. $\square$

Definition 63. Let X and Y be metric spaces, $E \subset X$ any subset, and $s > 0$. For any mapping $f : E \to Y$, we define the *lower s-density of f* as

$$\Theta_*^s(f,E,x) = \liminf_{r\to 0^+} \frac{\mathcal{H}_\infty^s(f(B(x,r)\cap E))}{\omega_s r^s}.$$

Remark 64. It is a routine exercise to show that we can replace open balls by closed balls in the definition of the lower density i.e.,

$$\Theta_*^s(f,E,x) = \liminf_{r\to 0^+} \frac{\mathcal{H}_\infty^s(f(\bar{B}(x,r)\cap E))}{\omega_s r^s}.$$

Remark 65. Note that if f is Lipschitz, then $\Theta_*^s(f,E,x) \leq (\operatorname{Lip} f)^s$.

Remark 66. In the case when $X = \mathbb{R}^k$, $s = n$, and Y is any metric space, the lower (and upper) n-density of f will show up in Chapter 6. These densities were introduced in [27] and it played an important role in the implicit function theorem for Lipschitz mappings into metric spaces.

Definition 67. We say that a Borel measure μ on X is doubling if $0 < \mu(B(x,r)) < \infty$ for all $x \in X$ and $r > 0$, and if there is a constant $C > 0$ such that $\mu(B(x,2r)) \leq C\mu(B(x,r))$ for all $x \in X$ and $r > 0$.

The next definition provides a particularly important instance of a doubling measure.

Definition 68. We say that the Hausdorff measure $\mathcal{H}^s$, $s > 0$, on X is *Ahlfors regular*, if there are constants $C_A, C_B > 0$ such that $C_A r^s \leq \mathcal{H}^s(B(x,r)) \leq C_B r^s$ for all $x \in X$ and all $r < \operatorname{diam} X$.

Definition 69. We say that a metric space is *metric doubling* if there is $M > 0$ such that every ball B can be covered by no more than M balls of half the radius.

Note that if a metric space is metric doubling, then bounded sets are totally bounded. Recall that a metric space is compact if and only it it is complete and totally bounded. Therefore we have

Lemma 70. *If X is metric doubling and complete, then X is boundedly compact.*

The following lemma is an easy exercise

Lemma 71. *If μ is a doubling measure on X, then X is metric doubling.*

Indeed, there cannot be too many points in B whose mutual distances are greater than or equal to $r/2$, where r is the radius of B.

The next result is the Vitali covering theorem for doubling measures, see [28, Theorem 1.6]

Lemma 72. *Let μ be a doubling measure on a metric space X and let $E \subset X$. If $\mathcal{F}$ is a family of closed balls centered at E such that for every $x \in E$*

$$\inf\{r > 0 : B(x,r) \in \mathcal{F}\} = 0,$$

then there is a countable subfamily $\{B_1, B_2, \ldots\} \subset \mathcal{F}$ of pairwise disjoint balls such that

$$\mu\Big(E \setminus \bigcup_{i=1}^{\infty} B_i\Big) = 0.$$

The next result is the Lebesgue differentiation theorem for doubling measures. It is a consequence of Lemma 72, see [28, Theorem 1.8]

Lemma 73. *If g is a locally integrable function on a metric space with a doubling measure μ, then*

$$\lim_{r\to 0} \fint_{B(x,r)} g\, d\mu = g(x) \quad \textit{for } \mu\textit{-almost all } x \in X. \tag{4.7}$$

Lemma 74. *Suppose the metric space X is metric doubling and $E \subset X$ is bounded. If $s, t \in [0, \infty)$, and $f : E \to Y$ is a mapping, then $\Phi^{s,t}(f, E) = 0$ if and only if $\Phi^{s,t}_\infty(f, E) = 0$.*

Proof. Since $\Phi^{s,t}_\infty \le \Phi^{s,t}$, one implication is obvious. It remains to show that if $\Phi^{s,t}_\infty(f, E) = 0$, then for any $\delta > 0$ we have $\Phi^{s,t}_\delta(f, E) = 0$. Since E is bounded and X is metric doubling, E can be split into a finite number of pieces, say $N(\delta)$ many, each of diameter less than δ.

Given $\varepsilon > 0$, let $E \subset \bigcup_{i=1}^\infty A_i$ be a covering such that

$$\sum_{i=1}^{\infty} \zeta^s(f(A_i))\zeta^t(A_i) < \frac{\varepsilon}{N(\delta)}.$$

By replacing A_i with $E \cap A_i$ we can further assume that $A_i \subset E$. Each of the sets A_i is a union of $N(\delta)$ sets $\{A_{ij}\}_{j=1}^{N(\delta)}$, each of diameter less that δ. Therefore,

$$\Phi^{s,t}_\delta(f, E) \le \sum_{i=1}^{\infty}\sum_{j=1}^{N(\delta)} \zeta^s(f(A_{ij}))\zeta^t(A_{ij}) \le N(\delta) \sum_{i=1}^{\infty} \zeta^s(f(A_i))\zeta^t(A_i) < \varepsilon.$$

□

Theorem 75. *Suppose $0 < t \le s < \infty$, the measure $\mathcal{H}^s$ is Ahlfors regular on a complete metric space X, $E \subset X$ is closed, and $f : E \to Y$ is Lipschitz. Then*

$$\int_Y^* \mathcal{H}^{s-t}(f^{-1}(y) \cap E)\, d\mathcal{H}^t(y) \le \frac{\omega_{s-t}\omega_t}{C_A} \int_E \Theta^t_*(f, E, x)\, d\mathcal{H}^s(x)\,. \tag{4.8}$$

where C_A is the constant from Definition 68.

Remark 76. The assumption that X is complete guarantees that X is boundedly compact (Lemma 70). Since E is closed, $\bar{B}(x,r)\cap E$ is compact. We need this assumption to prove measurability of $\Theta^t_*(f,E,\cdot)$. We do not know if the theorem is true for any $\mathcal{H}^s$-measurable set E, and without assuming that X is complete.

Proof. We can assume that E is bounded, because the general case will follow from the inequality applied to $E\cap\bar{B}(x_o,R)$ upon passing to the limit as $R\to\infty$. Note that in order to pass to the limit on the left hand side, we need to use Lemma 11.

The density function $\Theta^t_*(f,E,\cdot)$ is measurable. To see this it suffices to prove that the function $h_r(x)=\mathcal{H}^s_\infty(f(\bar{B}(x,r)\cap E))$ (see Remark 64) is Borel and this is true since the function is upper-semicontinuous meaning that $\limsup_{y\to x}h_r(y)\leq h_r(x)$. Indeed, under our assumptions, the set $\bar{B}(x,r)\cap E$ and its image are compact. We can approximate $\mathcal{H}^s_\infty(f(\bar{B}(x,r)\cap E))$ using an open covering $\{U_i\}_{i=1}^\infty$. If y is close to x, then $f(\bar{B}(y,r)\cap E)\subset\bigcup_{i=1}^\infty U_i$ and we can use the same open covering $\{U_i\}_{i=1}^\infty$ to get the upper estimate for the content $\mathcal{H}^s_\infty(f(\bar{B}(y,r)\cap E))$.

Since $\mathcal{H}^s(E)<\infty$ (E is bounded and $\mathcal{H}^s$ is Ahlfors regular), in view of Remark 65, the right hand side of (4.8) is finite.

According to Theorem 58, it suffices to prove that

$$\Phi^{t,s-t}(f,E)\leq\frac{\omega_{s-t}\omega_t}{C_A}\int_E\Theta^t_*(f,E,x)\,d\mathcal{H}^s(x)\,. \tag{4.9}$$

Let N be the set of points $x\in E$ for which (4.7) does not hold with $g=\Theta^t_*(f,E,\cdot)\chi_E$. Since $\mathcal{H}^s(N)=0$, Lemma 57 yields that $\Phi^{t,s-t}(f,N)=0$ and hence by Lemma 56,

$$\Phi^{t,s-t}(f,E)=\Phi^{t,s-t}(f,N)\,. \tag{4.10}$$

Given $\varepsilon>0$ and $\delta>0$, for each $x\in E\setminus N$, there is a sequence $r_{x,i}\to 0^+$, $B_{x,i}=\bar{B}(x,r_{x,i})$ such that

$$\frac{\mathcal{H}^t_\infty(f(E\cap B_{x,i}))}{\omega_t r^t_{x,i}}\leq\Theta^t_*(f,E,x)+\frac{\varepsilon}{2}\leq \fint_{B(x,r_{x,i})}\Theta^t_*(f,E,z)\chi_E(z)\,d\mathcal{H}^s(z)+\varepsilon.$$

Lemma 72 applied to the family $\{B_{x,i} : x \in E \setminus N, r_{x,i} < \delta/2\}$ gives pairwise disjoint balls B_i with diameters less than δ such that

$$\mathcal{H}^s\Big(E \setminus \bigcup_{i=1}^{\infty} B_i\Big) = \mathcal{H}^s\Big((E \setminus N) \setminus \bigcup_{i=1}^{\infty} B_i\Big) = 0.$$

Using a similar argument as in the proof of (4.10), one can easily show that

$$\Phi_\delta^{t,s-t}(f, E) = \Phi_\delta^{t,s-t}\Big(f, E \cap \bigcup_{i=1}^{\infty} B_i\Big).$$

Therefore, Lemma 62 yields

$$\begin{aligned}
&\Phi_\delta^{t,s-t}(f, E) = \Phi_\delta^{t,s-t}\Big(f, E \cap \bigcup_{i=1}^{\infty} B_i\Big) = \tilde{\mathcal{H}}_\delta^{t,s-t}\Big(f, E \cap \bigcup_{i=1}^{\infty} B_i\Big) \\
&\leq \sum_{i=1}^{\infty} \zeta^{s-t}(E \cap B_i)\mathcal{H}_\infty^t(f(E \cap B_i)) \\
&\leq \sum_{i=1}^{\infty} \frac{\omega_{s-t}}{2^{s-t}}(2r_i)^{s-t}\omega_t r_i^t\Big(\fint_{B(x_i,r_i)} \Theta_*^t(f, E, z)\chi_E(z)\, d\mathcal{H}^s(z) + \varepsilon\Big) \\
&\leq \frac{\omega_{s-t}\omega_t}{C_A}\Big(\int_E \Theta_*^t(f, E, z)\, d\mathcal{H}^s(z) + \varepsilon\Big)
\end{aligned}$$

and the result follows by letting $\delta \to 0^+$ and then $\varepsilon \to 0^+$. $\square$

It was proved in [27, Proposition 5.2] that if $f : E \to \mathbb{R}^m$ is a Lipschitz continuous map defined on a measurable set $E \subset \mathbb{R}^n$, $n \geq m$, then $\Theta_*^m(f, E, x) = |J^m f|(x)$, where

$$|J^m f|(x) = \sqrt{\det(Df)(Df)^T} \quad \text{is the Jacobian.}$$

This and the above result gives

Corollary 77. *If $f : E \to \mathbb{R}^m$ is a Lipschitz map defined on a measurable set $E \subset \mathbb{R}^n$, $n \geq m$, then*

$$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(f^{-1}(y) \cap E)\, d\mathcal{H}^m(y) \leq \frac{\omega_{n-m}\omega_m}{\omega_n} \int_E |J^m f|(x)\, d\mathcal{H}^n(x).$$

Remark 78. The celebrated coarea formula, Lemma 104, states that under the above assumptions

$$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(f^{-1}(y) \cap E)\, d\mathcal{H}^m(y) = \int_E |J^m f|(x)\, d\mathcal{H}^n(x).$$

Since we obtained the Corollary 77 as a consequence of rather general results valid in metric spaces, it is not surprising that the result is not as sharp as the coarea formula. On the other hand a localized version of Theorem 52 would suggest a much weaker inequality with $|Df|^m$ instead of $|J^m f|$ since $|Df|$ can be regarded as a local Lipschitz constant of f. This shows that Theorem 75 and hence also Theorem 58 are substantial improvements of the coarea inequality.

5.0 Metric differentiability and the area formula

The theorem of metric differentiability was developed by Kirchheim in his 1994 paper [34], and can be considered a classic now. However, the success of the many applications of this notion in this thesis relies heavily on the link between the metric derivative and the so-called componentwise derivative, which makes geometric applications easy (e.g., see Proposition 121). The componentwise derivative has been previously investigated in [25, 26, 27], but without connection to the metric derivative.

If $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is (Fréchet) differentiable at $x \in \mathbb{R}^n$, then

$$\lim_{y\to x} \frac{f(y) - f(x) - Df(x)(y-x)}{|y-x|} = 0,$$

and it follows from the triangle inequality that

$$\lim_{y\to x} \frac{|f(y) - f(x)| - \sigma_x(y-x)}{|y-x|} = 0,$$

where $\sigma_x(v) = |Df(x)v|$ is a seminorm on $\mathbb{R}^n$. Recall that $\sigma\colon \mathbb{R}^n \to [0,\infty)$ is a seminorm if $\sigma(\lambda v) = |\lambda|\sigma(v)$ and $\sigma(v+w) \leq \sigma(v) + \sigma(w)$ for all $\lambda \in \mathbb{R}$ and $v, w \in \mathbb{R}^n$. Thus a seminorm is like a norm, but it may vanish on a non-trivial linear subspace of $\mathbb{R}^n$.

Denote by ℓ^∞_m the space $\mathbb{R}^m$ with the norm $\|v\|_\infty = \max_{1\leq i\leq m} |v_i|$, $v = (v_1, \ldots, v_m)$. Again, if f is as above, but we regard $\mathbb{R}^m$ as ℓ^∞_m, then the triangle inequality yields

$$\lim_{y\to x} \frac{\|f(y) - f(x)\|_\infty - \sigma^\infty_x(y-x)}{|y-x|} = 0, \quad \text{where} \quad \sigma^\infty_x(v) = \|Df(x)v\|_\infty. \tag{5.1}$$

The above observations motivate the following definition due to Kirchheim [34].

Definition 79. Let $f : \Omega \to X$ be a map between an open set $\Omega \subset \mathbb{R}^n$ and a metric space (X, d). We say that f is *metrically differentiable* at $x \in \Omega$, if there is a seminorm σ_x on $\mathbb{R}^n$ such that

$$\lim_{y\to x} \frac{d(f(y), f(x)) - \sigma_x(y-x)}{|y-x|} = 0.$$

If f is metrically differentiable at x, then the seminorm σ_x is unique (easy exercise) and we denote it by $\mathrm{md}\,(f,x)$, i.e. ,

$$\lim_{y\to x} \frac{d(f(y), f(x)) - \mathrm{md}\,(f,x)(y-x)}{|y-x|} = 0.$$

The seminorm $\operatorname{md}(f,x)$ is called the *metric derivative* of f at x.

It follows that for any vector $v \in \mathbb{R}^n$,

$$\lim_{t\to 0} \frac{d(f(x+tv), f(x))}{|t|} = \operatorname{md}(f,x)(v). \tag{5.2}$$

Thus the "directional speed" of f exists at x in every direction v, and it defines as a function of v, a seminorm on $\mathbb{R}^n$. If f is L-Lipschitz, we also get the estimates

$$\operatorname{md}(f,x)(v) \le L|v| \quad \text{and} \quad |\operatorname{md}(f,x)(v) - \operatorname{md}(f,x)(w)| \le L|v-w|. \tag{5.3}$$

One, however, needs to be aware that the metric differentiability is much weaker than the Fréchet differentiability as the next example shows.

Example 80. $f:\mathbb{R}\to\mathbb{R}$, $f(x)=|x|$, is metrically differentiable at 0 and $\operatorname{md}(f,0)(v)=|v|$.

Example 81. By Rademacher's theorem, Lipschitz continuous maps $f:\mathbb{R}^n \supset \Omega \to \mathbb{R}^m$ are differentiable a.e. and hence (5.1) implies that Lipschitz continuous maps

$$f=(f_1,f_2,\ldots,f_m):\mathbb{R}^n \supset \Omega \to \ell^\infty_m$$

are metrically differentiable a.e. with

$$\operatorname{md}(f,x)(v) = \|Df(x)v\|_\infty = \max_{1\le i\le m} |\nabla f_i(x)\cdot v|.$$

This example is relevant for the case of mappings into arbitrary metric spaces, because any separable metric space admits an isometric embedding into ℓ^∞, see Theorem 37.

The next result, a far reaching generalization of Example 81, and its corollary are due to Kirchheim [34] (see also [1]). They are known as the *Kirchheim-Rademacher theorem.*

Theorem 82. *Let $f:\mathbb{R}^n\supset\Omega\to\ell^\infty$, $f=(f_1,f_2,\ldots)$ be a Lipschitz mapping. Then f is metrically differentiable a.e., and*

$$\operatorname{md}(f,x)(v) = \sup_{i\in\mathbb{N}} |\nabla f_i(x)\cdot v| \quad \textit{for almost all } x\in\Omega \textit{ and all } v\in\mathbb{R}^n. \tag{5.4}$$

Remark 83. By Rademacher's theorem, Theorem 34, each component f_i is differentiable a.e. Since the union of countably many sets of measure zero has measure zero, there is a Borel set $N \subset \Omega$ of measure zero, $|N| = 0$, such that for each $i \in \mathbb{N}$ and all $x \in D := \Omega \setminus N$, f_i is differentiable at x. Therefore, the expression on the right hand side of (5.4) is well defined for all almost all $x \in \Omega$ (namely for all $x \in D$) and for all $v \in \mathbb{R}^n$.

Remark 84. Note that for every v, the function $x \mapsto \mathrm{md}\,(f, x)(v)$ is measurable and for almost every x, the function $v \mapsto \mathrm{md}\,(f, x)(v)$ is continuous. Regarding measurability of $x \mapsto \mathrm{md}\,(f, x)(v)$, by restricting the metric derivative to the Borel set $D = \Omega \setminus N$, where the right hand side of (5.4) is well defined (cf. Remark 83), we see that for every $v \in \mathbb{R}^n$, the function $D \ni x \mapsto \mathrm{md}\,(f, x)(v)$ is Borel. While, in Theorem 82 we assume that the mapping is into ℓ^∞, the remark applies to Lipschitz mappings $f : \Omega \to X$ into arbitrary metric spaces, because the image $f(\Omega)$ is separable and hence can be isometrically embedded into ℓ^∞.

Corollary 85. *Let $f : \mathbb{R}^n \supset \Omega \to X$ be a Lipschitz map into a metric space. Then f is metrically differentiable a.e.*

Proof. Although X is not required to be separable, the subset $f(\Omega) \subset X$ is separable and hence it admits an isometric embedding into ℓ^∞, by Lemma 11. Then, the corollary follows from Theorem 82. $\square$

Before we prove Theorem 82 let us introduce some terminology and explain why Theorem 82 is far from being obvious.

Definition 86. Let $f : \mathbb{R}^n \supset \Omega \to \ell^\infty$, $f = (f_1, f_2, \ldots)$ be a Lipschitz map and let the set $N \subset \Omega$ be defined as in Remark 83. For $x \in \Omega \setminus N$, the *componentwise derivative* of f at x is a linear map $Df(x) : \mathbb{R}^n \to \ell^\infty$ defined by

$$Df(x)v = \big(\nabla f_1(x) \cdot v, \nabla f_2(x) \cdot v, \ldots\big).$$

Indeed, if f is L-Lipschitz, then each component f_i is L-Lipschitz, so for $x \in \Omega \setminus N$ and $v \in \mathbb{R}^n$, the estimate

$$\|Df(x)v\|_\infty = \sup_{i \in \mathbb{N}} |\nabla f_i(x) \cdot v| \leq L|v|$$

proves that $Df(x)$ maps $\mathbb{R}^n$ (linearly) into ℓ^∞.

Lemma 87. *If a Lipschitz map $f : \mathbb{R}^n \supset \Omega \to \ell^\infty$ is Fréchet differentiable at $x \in \Omega$, then the componentwise derivative $Df(x)$ is well-defined and equals the Fréchet derivative. Moreover,*

$$\mathrm{md}\,(f,x)(v) = \|Df(x)(v)\|_\infty = \sup_{i\in\mathbb{N}} |\nabla f_i(x)\cdot v|\,.$$

Proof. Let a linear map $L = (L_1, L_2, \ldots) : \mathbb{R}^n \to \ell^\infty$ be the Fréchet derivative of f at x. Then

$$\frac{\|f(y)-f(x)-L(y-x)\|_\infty}{|y-x|} = \sup_{i\in\mathbb{N}} \frac{|f_i(y)-f_i(x)-L_i(y-x)|}{|y-x|} \to 0 \quad \text{as } y \to x.$$

It follows that for each $i \in \mathbb{N}$, f_i is differentiable at x and $\nabla f_i(x) = L_i$. Therefore, $L = Df(x)$, the componentwise derivative of f at x.

The triangle inequality yields

$$\frac{\|f(y)-f(x)\|_\infty - \|Df(x)(y-x)\|_\infty}{|y-x|} \to 0 \quad \text{as } y \to x$$

and hence $\mathrm{md}\,(f,x)(v) = \|Df(x)(v)\|_\infty$ is the metric derivative of f at x. □

Lemma 87 is a generalization of Example 81 and gives metric derivative in terms of Fréchet derivative. However, this is by no means a path to a proof of Theorem 82. In fact, there are Lipschitz mappings $f : \mathbb{R}^n \supset \Omega \to \ell^\infty$ that are *nowhere* Fréchet differentiable. The next example is well known.

Proposition 88. *The Lipschitz mapping $f : (0,1) \to \ell^\infty$ defined by*

$$f(x) = (f_1(x), f_2(x), \ldots), \quad f_i(x) = \frac{\sin(ix)}{i},$$

is nowhere Fréchet differentiable.

Proof. First, note that f maps $(0,1)$ into a closed subspace $c_0 \subset \ell^\infty$ and that f is indeed, Lipschitz continuous:

$$\|f(x) - f(y)\|_\infty = \sup_{i \in \mathbb{N}} \left| \frac{\sin(ix) - \sin(iy)}{i} \right| \leq |x - y|.$$

Suppose to the contrary that f is Fréchet differentable at $x \in (0,1)$. Then by Lemma 87 the Fréchet derivative equals the componentwise derivative

$$Df(x) : \mathbb{R} \to \ell^\infty, \quad Df(x)t = (t\cos x, t\cos(2x), t\cos(3x), \dots)$$

and the Fréchet differentiability yields

$$\left\| \frac{f(x+t) - f(x) - Df(x)t}{t} \right\|_\infty = \left\| \frac{f(x+t) - f(x)}{t} - Df(x)1 \right\|_\infty \to 0 \quad \text{as } t \to 0.$$

This is however, impossible because $(f(x+t) - f(x))/t \in c_0$ while

$$Df(x)1 = (\cos x, \cos(2x), \cos(3x), \dots) \in \ell^\infty \setminus c_0$$

and an element of $\ell^\infty \setminus c_0$ cannot be approximated by elements of c_0 in the ℓ^∞ norm. $\square$

5.1 Proof of Kirchheim-Rademacher theorem

Proof of Theorem 82. Since the result is local in nature, we may assume that $\Omega = \mathbb{R}^n$. This will slightly simplify our notation. Let $N \subset \mathbb{R}^n$ be a set of measure zero as in Remark 83 and let $Df(x) : \mathbb{R}^n \to \ell^\infty$, for $x \in \mathbb{R}^n \setminus N$, be the componentwise derivative. Then for any $v \in \mathbb{R}^n$ we have

$$\|Df(x)v\|_\infty \leq \liminf_{t\to 0} \left\| \frac{f(x+tv)-f(x)}{t} \right\|_\infty . \tag{5.5}$$

Indeed, if $i \in \mathbb{N}$, then

$$|\nabla f_i(x) \cdot v| = \lim_{t\to 0} \left| \frac{f_i(x+tv)-f_i(x)}{t} \right| \leq \liminf_{t\to 0} \left\| \frac{f(x+tv)-f(x)}{t} \right\|_\infty$$

and (5.5) follows upon taking the supremum over $i \in \mathbb{N}$.

We will prove now that in fact, we have a stronger equality in (5.5). Namely, for almost all $x \in \mathbb{R}^n$ and all $v \in \mathbb{R}^n$

$$\|Df(x)v\|_\infty = \lim_{t\to 0} \left\| \frac{f(x+tv)-f(x)}{t} \right\|_\infty .$$

Fix $0 \neq v \in \mathbb{R}^n$. Assume that the line $\ell = \{x_o + tv : t \in \mathbb{R}\}$ intersects N along a set of length zero. Then for almost all $z \in \ell$ (namely for all $z \in \ell \setminus N$) and all $i \in \mathbb{N}$, the directional derivative satisfies $D_v f_i(z) = \nabla f_i(x) \cdot v$. Since functions $f_i|_\ell$ are Lipschitz continuous, it follows that for all $x \in \ell$, all $t \in \mathbb{R}$ and all $i \in \mathbb{N}$,

$$f_i(x+tv) - f_i(x) = \int_0^t \frac{d}{d\tau} f_i(x+\tau v)\, d\tau = \int_0^t \nabla f_i(x+\tau v) \cdot v\, d\tau. \tag{5.6}$$

Let W be the union of all lines ℓ that intersect N along a set of length zero. By Fubini's theorem $|\mathbb{R}^n \setminus W| = 0$. Fix $x \in W$ and $t \in \mathbb{R}$. For any $\varepsilon > 0$ there is $i \in \mathbb{N}$ such that

$$\begin{aligned} \|f(x+tv)-f(x)\|_\infty - \varepsilon \leq |f_i(x+tv)-f_i(x)| &= \left| \int_0^t \nabla f_i(x+\tau v)\cdot v\, d\tau \right| \\ &\leq \int_0^t \|Df(x+\tau v)v\|_\infty\, d\tau. \end{aligned}$$

Since this inequality is true for any $\varepsilon > 0$, we have that for all $x \in W$,

$$\|f(x+tv)-f(x)\|_\infty \leq \int_0^t \|Df(x+\tau v)v\|_\infty\, d\tau \quad \text{for all } t \in \mathbb{R}. \tag{5.7}$$

All lines $\ell = \{x_o + tv : t \in \mathbb{R}\} \subset W$ have the following two properties:

(a) The following function is measurable and bounded:

$$\tau \mapsto \|Df(x_o+\tau v)v\|_\infty = \sup_{i\in\mathbb{N}} \left| \frac{d}{d\tau} f_i(x_o+\tau v)\right|$$

(b) For all $s\in\mathbb{R}$, points $x = x_o+sv\in\ell$ satisfy (5.7) i.e.,

$$\|f((x_o+sv)+tv)-f(x_o+sv)\|_\infty \le \int_s^{t+s} \|Df((x_o+\tau v)v\|_\infty\, d\tau.$$

Now (a) and Lebesgue's differentation theorem imply that for almost all $s\in\mathbb{R}$,

$$\lim_{t\to 0}\frac{1}{|t|}\int_s^{t+s} \|Df(x_o+\tau v)v\|_\infty\, d\tau = \|Df(x_o+sv)v\|_\infty,$$

which together with (b) yield that for almost all $s\in\mathbb{R}$,

$$\limsup_{t\to 0}\left\|\frac{f((x_o+sv)+tv)-f(x_o+sv)}{t}\right\|_\infty \le \|Df(x_o+sv)v\|_\infty\,.$$

Since this is true for almost all points $x_o+sv\in\ell$ on all lines $\ell\subset W$, we conclude that there is a set $N_v\subset\mathbb{R}^n$ of measure zero $|N_v|=0$ such that for all $x\in\mathbb{R}^n\setminus N_v$ we have

$$\limsup_{t\to 0}\left\|\frac{f(x+tv)-f(x)}{t}\right\|_\infty \le \|Df(x)v\|_\infty. \tag{5.8}$$

For each $0\neq v\in\mathbb{R}^n$ we have a different exceptional set N_v. Let $\{v_i\}_{i=1}^\infty\subset\mathbb{S}^{n-1}$ be countable and dense. Let $\tilde{N}=\bigcup_{i=1}^\infty N_{v_i}$. Clearly, $|\tilde{N}|=0$. We will prove that for all $x\in\mathbb{R}^n\setminus\tilde{N}$, (5.8) is true for all $v\in\mathbb{R}^n$.

Let $x\in\mathbb{R}^n\setminus\tilde{N}$. Then (5.8) is true for all $v=v_i$. Since both sides of (5.8) are 1-homogeneous with respect to v, (5.8) is also true for $v=\lambda v_i$, $\lambda>0$. Note that the set $V:=\{\lambda v_i\,:\,\lambda>0,\ i\in\mathbb{N}\}\subset\mathbb{R}^n$ is dense. It is easy to check that both sides of (5.8) define functions that are Lipschitz continuous in v. So the fact that inequality (5.8) between Lipschitz functions is valid on a dense subset $V\subset\mathbb{R}^n$, implies that it is true for all $v\in\mathbb{R}^n$. Inequalities (5.8) and (5.5) yield that for almost all $x\in\mathbb{R}^n$ (namely for all $x\in\mathbb{R}^n\setminus(N\cup\tilde{N})$)

$$\|Df(x)v\|_\infty = \lim_{t\to 0}\left\|\frac{f(x+tv)-f(x)}{t}\right\|_\infty \quad \text{for all } v\in\mathbb{R}^n. \tag{5.9}$$

We now prove a stronger fact that $\mathrm{md}\,(f,x)(v) = \|Df(x)v\|_\infty$ is the metric derivative of f for all $x \in \mathbb{R}^n \setminus (N \cup \tilde{N})$. It is easy to see that $v \mapsto \|Df(x)v\|_\infty$ is a seminorm and it remains to show that

$$\lim_{t\to 0} \sup_{|v|=1} \left| \left\| \frac{f(x+tv)-f(x)}{t} \right\|_\infty - \|Df(x)v\|_\infty \right| = 0. \tag{5.10}$$

Let as before, $\{v_i\}_{i=1}^\infty \subset \mathbb{S}^{n-1}$ be a dense subset. Given $\varepsilon > 0$, there is $p \in \mathbb{N}$ such that for every $v \in \mathbb{S}^{n-1}$ there is $k \in \{1,2,\ldots,p\}$ such that

$$|v - v_k| < \frac{\varepsilon}{4L}, \tag{5.11}$$

where L is the Lipschitz constant of f. It follows from (5.9) that there is $\delta > 0$ such that for all $0 < |t| < \delta$

$$\sup_{1\le i\le p} \left| \left\| \frac{f(x+tv_i)-f(x)}{t} \right\|_\infty - \|Df(x)v_i\|_\infty \right| < \frac{\varepsilon}{2}.$$

Using an elementary inequality

$$\big|\|a\| - \|b\|\big| \le \big|\|a_k\| - \|b_k\|\big| + \|a - a_k\| + \|b - b_k\|,$$

for any $0 < |t| < \delta$, any $v \in \mathbb{S}^{n-1}$ and v_k satisfying (5.11), we have

$$\begin{aligned}
&\left| \left\| \frac{f(x+tv)-f(x)}{t} \right\|_\infty - \|Df(x)v\|_\infty \right| \le \left| \left\| \frac{f(x+tv_k)-f(x)}{t} \right\|_\infty - \|Df(x)v_k\|_\infty \right| \\
&+ \left\| \frac{f(x+tv)-f(x+tv_k)}{t} \right\|_\infty + \|Df(x)(v-v_k)\|_\infty \\
&\le \frac{\varepsilon}{2} + L|v-v_k| + L|v-v_k| < \varepsilon
\end{aligned}$$

and (5.10) follows. The proof is complete. $\square$

Observe that $C(\mathbb{S}^{n-1})$ and $C(\mathbb{S}^{n-1} \times [0,1])$ are separable metric spaces with metrics

$$d_\infty(g,h) = \sup_{|v|=1} |g(v) - h(v)| \quad \text{and} \quad \bar{d}_\infty(g,h) = \sup_{|v|=1} \sup_{0\le t\le 1} |g(v,t) - h(v,t)|$$

respectively. Separability easily follows from the Stone-Weierstrass theorem. Note also that the restriction of continuous functions on $\mathbb{S}^{n-1}\times[0,1]$ to $\mathbb{S}^{n-1}\times\{0\} \simeq \mathbb{S}^{n-1}$ yields a continuous (surjective) map

$$\pi : C(\mathbb{S}^{n-1} \times [0,1]) \to C(\mathbb{S}^{n-1}).$$

Let $D \subset \Omega$ be the Borel set of points where $f\colon \Omega \to X$ is metrically differentiable, and $|\Omega \setminus D| = 0$. Consider the map

$$\Phi_f : D \to C(\mathbb{S}^{n-1}), \quad \Phi_f(x)(v) = \mathrm{md}\,(f,x)(v),\ |v| = 1.$$

The next lemma provides an elementary, but useful estimate for the continuity of the metric derivative.

Lemma 89. *If $f : \mathbb{R}^n \supset \Omega \to X$ is L-Lipschitz and metrically differentiable at $x, y \in \Omega$, then for any $v, w \in \mathbb{R}^n$ we have*

$$|\mathrm{md}\,(f,x)(v) - \mathrm{md}\,(f,y)(w)| \le L|v-w| + \min\{|v|,|w|\}\, d_\infty\big(\Phi_f(x), \Phi_f(y)\big). \tag{5.12}$$

Proof. If $w = 0$ (or similarly if $v = 0$), (5.3) yields

$$|\mathrm{md}\,(f,x)(v) - \mathrm{md}\,(f,y)(w)| = \mathrm{md}\,(f,x)(v) \le L|v| = L|v-w|.$$

Thus we may assume that $v, w \neq 0$ and that $0 < |w| \le |v|$. Again (5.3) gives

$$\begin{aligned}
&|\mathrm{md}\,(f,x)(v) - \mathrm{md}\,(f,y)(w)| \\
&\le |\mathrm{md}\,(f,x)(v) - \mathrm{md}\,(f,x)(w)| + |w| \left| \mathrm{md}\,(f,x)\left(\frac{w}{|w|}\right) - \mathrm{md}\,(f,y)\left(\frac{w}{|w|}\right) \right| \\
&\le L|v-w| + |w| d_\infty(\Phi_f(x), \Phi_f(y)).
\end{aligned}$$

$\square$

Consider now the map

$$\Psi_f : D \to C(\mathbb{S}^{n-1} \times [0,1]), \quad \Psi_f(x)(v,t) = \begin{cases} \frac{d(f(x), f(x+tv))}{t} & \text{if } 0 < t \le 1 \\ \mathrm{md}\,(f,x)(v), & \text{if } t = 0. \end{cases}$$

Note that continuity of $\Psi_f(x) : \mathbb{S}^{n-1} \times [0,1] \to \mathbb{R}$ when $x \in D$ follows from the definition of metric differentiability.

Lemma 90. $\Psi_f : D \to C(\mathbb{S}^{n-1} \times [0,1])$ *is Borel measurable.*

Since $\pi : C(\mathbb{S}^{n-1} \times [0,1]) \to C(\mathbb{S}^{n-1})$ is continuous, and $\Phi_f = \pi \circ \Psi_f$, we immediately obtain

Corollary 91. $\Phi_f : D \to C(\mathbb{S}^{n-1})$ *is Borel measurable.*

Proof of Lemma 90. We need to prove that the preimage of any open set is Borel. To this end, it suffices to show that the preimage of any closed ball is Borel. Fix arbitrary $g \in C(\mathbb{S}^{n-1} \times [0,1])$ and $r > 0$. We need to show that the set $\Psi_f^{-1}(\bar{B}(g,r)) \subset D$ is Borel. We will show that in fact, this set is the intersection of a closed subset of Ω with D.

Let $\{v_i\}_{i=1}^{\infty} \subset \mathbb{S}^{n-1}$ and $\{t_j\}_{j=1}^{\infty} \subset (0,1]$ be countable and dense. Note that the sets

$$E_{ij} = \left\{ x \in \Omega : \left| \frac{d(f(x), f(x + t_j v_i))}{t_j} - g(v_i, t_j) \right| \le r \right\}$$

are closed subsets of Ω in the topology of Ω inherited from $\mathbb{R}^n$. We have

$$\begin{aligned}
\Psi_f^{-1}(\bar{B}(g,r)) &= \{x \in D : \bar{d}_\infty(\Psi_f(x), g) \le r\} \\
&= \{x \in D : |\Psi_f(x)(v,t) - g(v,t)| \le r \text{ for all } |v| = 1 \text{ and } 0 \le t \le 1\} \\
&= \{x \in D : |\Psi_f(x)(v_i,t_j) - g(v_i,t_j)| \le r \text{ for all } i,j \in \mathbb{N}\} \\
&= D \cap \bigcap_{i,j=1}^{\infty} E_{ij}.
\end{aligned}$$

□

Theorem 92. *Let $f : \mathbb{R}^n \supset \Omega \to X$ be Lipschitz continuous. Then for any $\varepsilon > 0$ there is a set $F_\varepsilon \subset \Omega$ which is closed as a subset of $\mathbb{R}^n$, such that $|\Omega \setminus F_\varepsilon| < \varepsilon$ and*

(a) $\Psi_f : F_\varepsilon \to C(\mathbb{S}^{n-1} \times [0,1])$ *is continuous.*

(b) $\operatorname{md}(f,\cdot)(\cdot) : F_\varepsilon \times \mathbb{R}^n \to \mathbb{R}$ *is continuous.*

(c) On every compact subset $K \subset F_\varepsilon$ we have

$$\lim_{\substack{|x-y|\to 0 \\ x\in K,\ y\in\mathbb{R}^n}} \frac{|d(f(x), f(y)) - \operatorname{md}(f,x)(y-x)|}{|y-x|} = 0. \tag{5.13}$$

Remark 93. Meaning of (5.13) is that $\forall\, \varepsilon > 0\ \exists\, \delta > 0\ \forall\, x \in K\ \forall\, y \in \mathbb{R}^n$

$$0 < |x-y| < \delta \implies \frac{|d(f(x), f(y)) - \operatorname{md}(f,x)(y-x)|}{|y-x|} < \varepsilon.$$

Remark 94. Part (b) can be regarded as a version of the Scorza-Dragoni theorem (cf. [7, Theorem 3.8]).

Proof. From Lemma 90 and Lusin's Theorem (Theorem 35), there is a set $F_\varepsilon \subset D$, which is closed as a subset of $\mathbb{R}^n$ such that $|D \setminus F_\varepsilon| < \varepsilon$ and

$$\Psi_f : F_\varepsilon \to C(\mathbb{S}^{n-1} \times [0,1]) \quad \text{and} \quad \Phi_f = \pi \circ \Psi_f : F_\varepsilon \to C(\mathbb{S}^{n-1})$$

are continuous. This proves (a). Now we show that (a) implies (b) and (c). It is easy to see that (b) follows from continuity of $\Phi_f : F_\varepsilon \to C(\mathbb{S}^{n-1})$ and from (5.12).

To prove (c), let $x_k \in K$, $x_k \neq y_k \in \mathbb{R}^n$, $|x_k - y_k| \to 0$ as $k \to \infty$. We can write $y_k = x_k + t_k v_k$, $|v_k| = 1$, $t_k \to 0^+$ and we need to show that

$$\frac{d(f(x_k), f(x_k + t_k v_k))}{t_k} - \operatorname{md}(f, x)(v_k) \to 0.$$

Suppose to the contrary that this is not the case, then after selecting subsequences, we may assume that $x_k \to x \in K$, $v_k \to v \in \mathbb{S}^{n-1}$, $0 < t_k \leq 1$ and

$$\left| \frac{d(f(x_k), f(x_k + t_k v_k))}{t_k} - \operatorname{md}(f, x)(v_k) \right| \geq \varepsilon \tag{5.14}$$

for some $\varepsilon > 0$ and all k. Note that (5.14) can be rewritten as

$$\Psi_f(x_k)(v_k, t_k) - \Psi_f(x_k)(v_k, 0)| \geq \varepsilon.$$

We have

$$\begin{aligned}
\varepsilon &\leq |\Psi_f(x_k)(v_k, t_k) - \Psi_f(x_k)(v_k, 0)| \leq |\Psi_f(x_k)(v_k, t_k) - \Psi_f(x)(v_k, t_k)| \\
&\quad + |\Psi_f(x)(v_k, t_k) - \Psi_f(x)(v_k, 0)| + |\Psi_f(x)(v_k, 0) - \Psi_f(x_k)(v_k, 0)| \\
&\quad \leq \bar{d}_\infty(\Psi_f(x_k), \Psi_f(x)) + |\Psi_f(x)(v_k, t_k) - \Psi_f(x)(v_k, 0)| + d_\infty(\Phi_f(x), \Phi_f(x_k)) \\
&\quad = A_k + B_k + C_k.
\end{aligned}$$

We used here the fact that

$$\Psi_f(x)(v_k, 0) = \Phi_f(x)(v_k) \quad \text{and} \quad \Psi_f(x_k)(v_k, 0) = \Phi_f(x_k)(v_k).$$

Clearly, $A_k, C_k \to 0$ as $k \to \infty$ by continuity of

$$\Psi_f : F_\varepsilon \supset K \to C(\mathbb{S}^{n-1} \times [0,1]) \quad \text{and} \quad \Phi_f : F_\varepsilon \supset K \to C(\mathbb{S}^{n-1}).$$

Finally, $B_k \to 0$ by uniform continuity of $\Psi_f(x) \colon \mathbb{S}^{n-1} \times [0,1] \to \mathbb{R}$. This however, contradicts the fact that $A_k + B_k + C_k \geq \varepsilon$. $\square$

By definition, the $\mathrm{md}\,(f,x)(y-x)$ approximates $d(f(y),f(x))$ when y is close to x. The next result asserts that in fact $d(f(y),f(z))$ is well approximated by $\mathrm{md}\,(f,x)(y-z)$ when both y and z are close to x. In other words, the pullback of the distance function locally resembles a seminorm.

Theorem 95. *If $f:\mathbb{R}^n \supset \Omega \to X$ is Lipschitz, then for almost all $x\in\Omega$, we have*

$$\lim_{\mathbb{R}^n\times\mathbb{R}^n\ni(y,z)\to(x,x)} \frac{d(f(y),f(z)) - \mathrm{md}\,(f,x)(y-z)}{|x-y|+|x-z|} = 0 \tag{5.15}$$

Remark 96. Sometimes, it is more convenient to write (5.15) as

$$d(f(y),f(z)) - \mathrm{md}\,(f,x)(y-z) = o(|x-y|+|x-z|).$$

We will use this notation in the proof.

Remark 97. Taking $y=x+tv$, $z=x+tw$, (5.15) yields that for any $v,w\in\mathbb{R}^n$

$$\lim_{t\to 0}\frac{d(f(x+tv),f(x+tw))}{|t|} = \mathrm{md}\,(f,x)(v-w)$$

which is a much stronger claim than the existence of the “directional speed” (5.2).

Proof. For $i=1,2,3,\ldots$ let $F_{1/i}\subset\Omega$ be a closed subset as in Theorem 92. Let $\tilde{F}_{1/i}$ be the set of density points of $F_{1/i}$. Since $F_{1/i}$ is closed, it follows that $\tilde{F}_{1/i}\subset F_{1/i}$. Let $E=\bigcup_{i=1}^{\infty}\tilde{F}_{1/i}$. Clearly, $|\Omega\setminus E|=0$. It suffices to show that (5.15) is true for all $x\in E$.

Let $x\in E$. Since y and z play a symmetric role in (5.15), it suffices to show that if

$$0<|x-y_k|\to 0 \quad\text{and}\quad |x-z_k|\le|x-y_k|, \tag{5.16}$$

then

$$\frac{d(f(y_k),f(z_k)) - \mathrm{md}\,(f,x)(y_k-z_k)}{|x-y_k|}\to 0$$

or that

$$d(f(y_k),f(z_k)) - \mathrm{md}\,(f,x)(y_k-z_k) = o(|x-y_k|). \tag{5.17}$$

Since $x\in E$, there is $i\in\mathbb{N}$ such that $x\in F_{1/i}$, and x is a density point of $F_{1/i}$. It easily follows from the definition of the density point that there is $\tilde{y}_k\in F_{1/i}$ such that

$$\frac{|y_k-\tilde{y}_k|}{|x-y_k|}\to 0 \text{ as } k\to\infty \quad\text{and}\quad |y_k-\tilde{y}_k|\le|x-y_k| \text{ for } k\ge k_o. \tag{5.18}$$

We have

$$
\begin{aligned}
&|d(f(y_k), f(z_k)) - \mathrm{md}\,(f,x)(y_k - z_k)| \le |d(f(y_k), f(z_k)) - d(f(\tilde{y}_k), f(z_k))| \\
&+ |d(f(\tilde{y}_k), f(z_k)) - \mathrm{md}\,(f,\tilde{y}_k)(\tilde{y}_k - z_k)| + |\mathrm{md}\,(f,\tilde{y}_k)(\tilde{y}_k - z_k) - \mathrm{md}\,(f,\tilde{y}_k)(y_k - z_k)| \\
&+ |\mathrm{md}\,(f,\tilde{y}_k)(y_k - z_k) - \mathrm{md}\,(f,x)(y_k - z_k)| = A_k + B_k + C_k + D_k.
\end{aligned}
$$

If follows from the triangle inequality and from (5.18) that

$$
A_k \le d(f(\tilde{y}_k, f(y_k)) \le L|\tilde{y}_k - y_k| = o(|x - y_k|).
$$

Inequality (5.3) yields

$$
C_k \le L|\tilde{y}_k - y_k| = o(|x - y_k|).
$$

Now, Theorem 92(c) implies that

$$
\frac{B_k}{|\tilde{y}_k - z_k|} \to 0 \quad \text{as } k \to \infty, \tag{5.19}
$$

because $\tilde{y}_k \in F_{1/i}$ as a convergent sequence, is contained in a compact subset of $F_{1/i}$. Inequalities (5.16) and (5.18) along with the triangle inequality imply that $|\tilde{y}_k - z_k| \le 3|x - y_k|$ for $k \ge k_o$ and hence $B_k = o(|x - y_k|)$, by (5.19).

It remains to estimate D_k. Since (5.16) yields $|y_k - z_k| \le 2|x - y_k|$, we have

$$
D_k \le |y_k - z_k|\, d_\infty(\Phi_f(\tilde{y}_k), \Phi_f(x)) = o(|x - y_k|),
$$

because $\Phi_f(\tilde{y}_k) \to \Phi_f(x)$ in $C(\mathbb{S}^{n-1})$ by continuity of Φ_f on $F_{1/i}$. This proves (5.17) and completes the proof of the theorem. $\square$

5.2 Rank of metric derivative and a Sard theorem

Unlike a norm, a seminorm may vanish on a non-trivial linear subspace of $\mathbb{R}^n$, $N_\sigma := \{v \in \mathbb{R}^n : \sigma(v) = 0\}$, and we define the *rank of a seminorm* σ on $\mathbb{R}^n$ as $\operatorname{rank}\sigma = n - \dim N_\sigma = \dim N_\sigma^\perp$. That is, it is the maximal dimension of a linear subspace on which σ is a norm.

Let $f : \mathbb{R}^n \supset \Omega \to \ell^\infty$, $f = (f_1, f_2, \ldots)$ be a Lipschitz mapping. By Rademacher's theorem and the fact that countable union of null sets is a null set, at a.e. x all $\nabla f_i(x)$ exist. We define the *componentwise derivative* of f to be the $\infty \times n$ matrix

$$Df(x) := \begin{bmatrix} \nabla f_1(x) \\ \nabla f_2(x) \\ \vdots \end{bmatrix}.$$

The next result is an easy exercise in linear algebra.

Lemma 98. *If $f = (f_1, f_2, \ldots) : \mathbb{R}^n \supset \Omega \to \ell^\infty$ is Lipschitz continuous, then for almost all $x \in \Omega$, the row rank of $Df(x)$ equals the column rank of $Df(x)$ and they equal* $\operatorname{rank}\operatorname{md}(f,x)$.

Now we use Theorem 95 to prove a result about covering of the image of a ball.

Proposition 99. *Let $f : \mathbb{R}^n \supset \Omega \to X$ be L-Lipschitz. Let*

$$E_k = \{x \in \Omega : \operatorname{rank}\operatorname{md}(f,x) = k\}, \quad 0 \le k \le n.$$

Then almost all points $x \in E_k$ have the following property: For every integer $m \ge 1$, there is $r_{x,m} > 0$ such that for all $0 < r < r_{x,m}$, $f(B(x,r))$ can be covered by m^k balls, each of radius $3\sqrt{k}Lr/m$, in the case of $k > 0$, and by one ball of radius r/m in the case $k = 0$.

Remark 100. This result is similar to [25, Lemma 2.7]. The approach in [25] uses componentwise differentiability instead of metric differentiability and as a result the proofs are different and more difficult.

Proof. Assume first that $k > 0$. Let $\tilde{E}_k$ be the set of all points $x \in E_k$ such that (5.15) holds. Clearly, $|E_k \setminus \tilde{E}_k| = 0$, and we will show that the property in the statement of the proposition is true for all $x \in \tilde{E}_k$. Fix $x \in \tilde{E}_k$, and let

$$N = \{v \in \mathbb{R}^n : \operatorname{md}(f, x)(v) = 0\} \quad \text{so} \quad \dim N = n - k. \tag{5.20}$$

By translating and rotating the coordinate system, we may assume that $x = 0$ and that

$$N^\perp = \operatorname{span}\{e_1, \dots, e_k\} \quad \text{and} \quad N = \operatorname{span}\{e_{k+1}, \dots, e_n\}.$$

Note that $B(x, r) = B(0, r) \subset [-r, r]^n = [-r, r]^k \times [-r, r]^{n-k}$. Given an integer $m \geq 1$, divide the cube $[-r, r]^k$ into a grid of m^k congruent cubes of edge length $2r/m$. Denote them by $\{Q_\nu\}_{\nu=1}^{m^k}$. Then

$$B(0, r) \subset \bigcup_{\nu=1}^{m^k} \big(Q_\nu \times [-r, r]^{n-k}\big), \qquad f(B(0, r)) \subset \bigcup_{\nu=1}^{m^k} f(Q_\nu \times [-r, r]^{n-k}).$$

So far, all this is true for any $r > 0$. Now it suffices to show that there is $r_{x,m} > 0$ such that for all $0 < r < r_{x,m}$ we have

$$\operatorname{diam}(f(Q_\nu \times [-r, r]^{n-k})) \leq 3\sqrt{k} L r/m. \tag{5.21}$$

In the case of $k = n$, this follows from $\operatorname{diam}(f(Q_\nu)) \leq L \operatorname{diam}(Q_\nu) = 2\sqrt{n} L r/m$. In the cases $0 < k < n$, (5.21) follows from (5.15). Since $x = 0$, (5.15) implies that there is $r_{x,m} > 0$ such that

$$|d(f(y), f(z)) - \operatorname{md}(f, 0)(y - z)| < \frac{\sqrt{k} L r}{m} \quad \text{for all } y, z \in [-r, r]^n,\ 0 < r < r_{x,m}. \tag{5.22}$$

In particular, it is true for $y, z \in Q_\nu \times [-r, r]^{n-k}$.

If $\pi : \mathbb{R}^n \to N^\perp = \mathbb{R}^k$ is the orthogonal projection onto the orthogonal complement of (5.20), then by triangle inequality and the fact that the Lipschitz constant L bounds $\operatorname{md}(f, \cdot)$,

$$\operatorname{md}(f, 0)(v) \leq \operatorname{md}(f, 0)(\pi(v)) + \underbrace{\operatorname{md}(f, 0)(v - \pi(v))}_{0} \leq L|\pi(v)|.$$

If $y, z \in Q_\nu \times [-r, r]^{n-k}$, then $\pi(y), \pi(z) \in Q_\nu$ so $|\pi(y-z)| \le 2\sqrt{k}r/m$, and hence

$$\operatorname{md}(f, 0)(y-z) \le 2\sqrt{k}Lr/m$$

which together with (5.22) gives $d(f(y), f(z)) < 3\sqrt{k}Lr/m$, and (5.21) follows.

Finally, if $k = 0$, then $\operatorname{md}(f, x) = 0$ and by definition of metric derivative there is $r_{x,m} > 0$ such that

$$d(f(y), f(x)) < \frac{|y-x|}{m} \quad \text{for all } y \in B(x, r), 0 < r < r_{x,m}.$$

But this shows that $f(B(x, r)) \subset B(f(x), r/m)$. □

Theorem 101. *Suppose $n \ge 1$ and $m \ge 0$ are integers, and $f \colon \mathbb{R}^{n+m} \supset \Omega \to X$ is Lipschitz. Let $E := \{x \in \Omega \colon \operatorname{rank} \operatorname{md}(f, x) < n\}$. Then*

$$\int_X \mathcal{H}^m \left(f^{-1}(y) \cap E\right) \, d\mathcal{H}^n(y) = 0 \, , \tag{5.23}$$

i.e. $\mathcal{H}^m (f^{-1}(y) \cap E) = 0$ for $\mathcal{H}^n$-a.e. $y \in X$.

Proof. By the coarea inequality (Theorem 58) it suffices to prove $\Phi^{n,m}(f, E) = 0$. For $0 \le k \le n-1$, let

$$E_k := \{x \in \Omega \colon \operatorname{rank} \operatorname{md}(f, x) = k\} \, .$$

Then $E = \bigcup_k E_k$. By countable subadditivity, it suffices to show $\Phi^{n,m}(f, A) = 0$ for every bounded $A \subset E_k$, $0 \le k \le n-1$.

Since $\Phi^{n,m}(f, N) = 0$ for any $\mathcal{L}^{n+m}$-null set, by Theorem 52, we may assume that *every* $x \in A$ has the following property: For every $0 < \delta \le 1$ and integer $N \ge 1$ there exists an $0 < r = r(x, N) \le \delta$ such that $f(B(x, r))$ has a covering by N^k balls of radius $\frac{Cr}{N}$ where $C = 3\sqrt{k}L$. Let's denote these balls by $B^i_j, j = 1, 2, \cdots, N^k$.

By Vitali's covering theorem, there is a covering of A by countably many such balls, say, $\{B(x_i, r_i)\}_i$, such that $B(x_i, r_i/5)$ are disjoint. Now, the family $A_{ij} := f^{-1}(B^i_j) \cap B(x_i, r_i)$, indexed by i and j, is a countable covering of A.

$$
\begin{aligned}
\Phi^{n,m}_\delta(f, A) &\lesssim \sum_i \sum_{j=1}^{N^k} (\operatorname{diam} f(A_{ij}))^n (\operatorname{diam} A_{ij})^m \\
&\lesssim \sum_i \sum_{j=1}^{N^k} (\frac{r_i}{N})^n (r_i)^m \\
&\leq \sum_i N^{k-n} r_i^{n+m} \\
&\lesssim N^{k-n} \sum_i (\frac{r_i}{5})^{n+m} \\
&\lesssim N^{k-n} (\mathcal{L}^{n+m}(A) + 1)\,.
\end{aligned}
$$

First let $N \to \infty$ to deduce $\Phi^{n,m}_\delta(f, A) = 0$. Then let $\delta \to 0^+$ to prove $\Phi^{n,m}(f, A) = 0$. $\square$

The special case of $m = 0$ in the previous theorem gives a Sard-type theorem.

Corollary 102. *Let $f : \mathbb{R}^n \supset \Omega \to X$ be Lipschitz. Let*

$$E = \{x \in \Omega :\ \operatorname{rank} \operatorname{md}(f, x) < n\},$$

Then $\mathcal{H}^n(f(E)) = 0$.

Proof. By (5.23),

$$\int_X \mathcal{H}^0\left(f^{-1}(y) \cap E\right)\, d\mathcal{H}^n(y) = 0\,.$$

But, $\mathcal{H}^0\left(f^{-1}(y) \cap E\right) = \operatorname{card}\{x \in E \colon f(x) = y\} \geq 1$ for every y in the $\mathcal{H}^n$-measurable set $f(E)$. So,

$$\mathcal{H}^n(f(E)) \leq \int_X \mathcal{H}^0\left(f^{-1}(y) \cap E\right)\, d\mathcal{H}^n(y) = 0\,.$$

$\square$

5.3 Metric area formula

The classical area formula states that if $f : \mathbb{R}^n \supset \Omega \to \mathbb{R}^m$ is Lipschitz, and $m \geq n$, and $g : \mathbb{R}^n \to [0, \infty]$ is measurable, then

$$\int_\Omega (g \circ f)|J_f|(x)\, d\mathcal{H}^n(x) = \int_{\mathbb{R}^m} g(y) N(f, \Omega, y)\, d\mathcal{H}^n(y). \tag{5.24}$$

Here $|J_f|(x) = \sqrt{\det(Df(x)^T Df(x))}$ is the Jacobian and the multiplicity function, a.k.a. the Banach indicatrix, is defined by

$$N(f, \Omega, y) := \mathcal{H}^0(f^{-1}(y) \cap \Omega) = \mathrm{card}\,(f^{-1}(y) \cap \Omega)\,.$$

One sees the simplest case of this formula in the calculation of (n-dimensional) surface area in calculus.

Kirchheim [34] proved an important generalization of the classical area formula to the setting of maps into arbitrary metric spaces using the notion of metric derivative.

We define the Jacobian of a seminorm $\sigma : \mathbb{R}^n \to [0, \infty)$ by

$$J_n(\sigma) = \frac{\omega_n}{\mathcal{H}^n(\{x :\, \sigma(x) \leq 1\}}\,.$$

Note that if σ is not a norm, i.e. if σ vanishes on a non-trivial linear subspace, then the set in the denominator is unbounded and it has infinite measure, so $J_n(\sigma) = 0$ in that case.

Theorem 103 (Kirchheim)**.** *Let $f : \Omega \to X$ be a Lipschitz mapping from an open set $\Omega \subset \mathbb{R}^n$ to a metric space X. Then*

$$\int_\Omega g(x) J_n(\mathrm{md}\,(f, x))\, d\mathcal{H}^n(x) = \int_{f(\Omega)} \Big(\sum_{x \in f^{-1}(y)} g(x) \Big)\, d\mathcal{H}^n(y) \tag{5.25}$$

for any Borel function $g : \Omega \to [0, \infty]$. In particular

$$\int_E g(f(x)) J_n(\mathrm{md}\,(f, x))\, d\mathcal{H}^n(x) = \int_X g(y)\, \mathcal{H}^0(E \cap f^{-1}(y))\, d\mathcal{H}^n(y) \tag{5.26}$$

for any Borel set $E \subset \Omega$ and any Borel function $g : X \to [0, \infty]$.

Notice that Corollary 102 gives the area formula on the part of the domain where $\operatorname{rank} \operatorname{md}(f, \cdot) < n$.

In the upcoming survey paper [17] we provide a different proof that fully utilizes our interpretation of the metric derivative via componentwise derivatives. The idea is as follows. Obviously one can assume without loss of generality that $X = \ell^\infty$. Then we prove that one recovers the area formula for $(f_1, f_2, \ldots)\colon \Omega \to \ell^\infty$ as the asymptote of the area formula for the truncations $(f_1, f_2, \ldots, f_j)\colon \Omega \to (\mathbb{R}^m, \|\cdot\|_{\ell^\infty_m})$ as $m \to \infty$. This is the hardest part to prove, because the area formula for the latter case is a simple modification of the classical area formula (5.24) above.

5.4 Metric coarea formula

Recall the classical coarea formula ([20, Theorem 3.10], [22, Theorem 3.2.11]).

Theorem 104. *If $f : \mathbb{R}^{n+m} \to \mathbb{R}^n$, $m \geq 0$, is Lipschitz and $E \subset \mathbb{R}^{n+m}$ is measurable, then*

$$\int_E |J^n f(x)| \, d\mathcal{H}^{n+m}(x) = \int_{\mathbb{R}^n} \mathcal{H}^m(f^{-1}(y) \cap E) \, d\mathcal{H}^n(y), \text{ where } |J^n f(x)| = \sqrt{\det(Df)(Df)^T}.$$

Again, using the notion of metric derivative, the coarea formula was subsequently and independently generalized by Karmanova [31] and Reichel [45] around 2008.

Theorem 105 (Metric coarea formula)**.** *Let $\Omega \subset \mathbb{R}^{n+m}$ (m can be zero) be open and X be an $\mathcal{H}^n$-σ-finite metric space, i.e. it is the union of countable many subsets with finite $\mathcal{H}^n$-measure. Suppose $f\colon \Omega \to X$ is Lipschitz. Then for any measurable $A \subset \Omega$*

$$\int_A C_n(\operatorname{md}(f, x)) \; d\mathcal{L}^{n+m}(x) = \int_X \mathcal{H}^m(f^{-1}(y) \cap A) \; d\mathcal{H}^n(y) \; . \tag{5.27}$$

Here *the coarea factor* $C_n(\sigma)$ is a geometrically defined quantity for seminorms with rank less than or equal to n. The coarea factor vanishes if $\operatorname{rank} \sigma < n$. Therefore, Theorem 101 gives the coarea formula on the part of the domain where $\operatorname{rank} \operatorname{md}(f, \cdot) < n$. In fact, our search for a coarea inequality that directly gave this implication was the motivation for our definition of the mapping content $\Phi^{s,t}$. This was acheived in Theorem 101. This reduces the proofs of the coarea formula to the set where the function is full-rank. On the latter set, the function has nicer geometry. In fact, in [17] we reduce the proof of the coarea formula to Fubini's theorem and the area formula for bi-Lipschitz maps.

The survey paper [17] will include a detailed exposition of the area and coarea formulas both in the Euclidean context and in metric context. In metric cases we provide new proofs for each.

6.0 Factorization through trees

This chapter is a collage of results that are unrelated in face value but their proofs use the same techniques, especially the notion of metric differentiability from Chapter 5. So, this chapter is logically dependent on Chapter 5. However, Chapters 5 and 6, together with the relevant sections from the preliminaries form a self-contained entity, which constitute the material for [18].

First, given a Lipschitz map f from a Euclidean cube into a metric space, we find several equivalent conditions for f to have a Lipschitz factorization through a metric tree. As an application we prove a recent conjecture of David and Schul. The techniques developed for the proof of the factorization result yield several other new and seemingly unrelated results. We prove that if f is a Lipschitz mapping from an open set in $\mathbb{R}^n$ onto a metric space X, then the topological dimension of X equals n if and only if X has positive n-dimensional Hausdorff measure. We also prove an area formula for length-preserving maps between metric spaces, which gives, in particular, a new formula for integration on countably rectifiable sets in the Heisenberg group.

In this Chapter (alone) we follow a convention that new results are denoted as a *Theorem* or a *Proposition*, while important known results are cited as a *Lemma* or a *Corollary*.

6.1 Main results and overview

Given metric spaces X, Y, Z, and a Lipschitz map $f : X \to Y$, we say that f *factors through* Z if there are Lipschitz mappings $\psi : X \to Z$ and $\phi : Z \to Y$ such that $f = \phi \circ \psi$.

Given a Lipschitz map $f \colon X \to Y$, our aim is to construct a space Z with a simple structure, along with a factorization $f = \phi \circ \psi$. In particular, we are interested in answering the question under what conditions, f factors through a metric tree (see Section 2.5 for the definition of a metric tree). This question was partially motivated by the recent works of Wenger and Young [50] and David and Schul [8]. The next result which is one of the main results of the chapter, provides several equivalent conditions for a factorization of a Lipschitz map through a metric tree.

Throughout the chapter n and m will stand for nonnegative integers.

Theorem 106. *If $f : Q_o = [0,1]^n \to X$, $n \geq 2$, is a Lipschitz map into a metric space, then the following conditions are equivalent:*

(a) f *factors through a metric tree.*

(b) $\operatorname{rank} \operatorname{md}(f,x) \leq 1$ *almost everywhere.*

(c) $\Theta^{*2}(f,x) = 0$ *almost everywhere.*

(d) $\Theta^2_*(f,x) = 0$ *almost everywhere.*

(e) $\mathcal{H}^{2,n-2}_\infty(f,Q_o) = 0$.

(f) $\hat{\mathcal{H}}^{2,n-2}_\infty(f,Q_o) = 0$.

Remark 107. In fact, in (a) we obtain quantitative estimates for the Lipschitz constants. Precisely, if f is L-Lipschitz, then we find a metric tree Z and maps $\psi\colon Q_o \to Z$ and $\phi\colon Z \to X$ such that ψ is L-Lipschitz, ϕ is 1-Lipschitz and $f = \phi \circ \psi$. The bounds follow from Lemma 139 and the fact that the cube is 1-quasiconvex.

Notation used in Theorem 106 will be introduced below after Theorem 108.

Equivalence of conditions (a) and (e) proves a recent conjecture of David and Schul [8, Conjecture 1.13]. They conjectured that if $f\colon Q_o = [0,1]^3 \to X$ satisfies $\mathcal{H}^{2,1}_\infty(f,Q_o) = 0$, then f factors through a metric tree.

Recently David and Schul [9], used our result (implication (e) $\Rightarrow$ (a)) to prove a quantitative part of Conjecture 1.13 from [8] which states that if the content $\mathcal{H}^{2,1}_\infty(f,Q_o)$ is small, then f is close to a mapping g that factors through a tree.

More precisely, they proved that for every $\varepsilon > 0$ there is $\delta = \delta(\varepsilon, m)$, $m \geq 1$, such that if $f : Q_o = [0,1]^{2+m} \to \ell^\infty$ satisfies $\mathcal{H}^{2,m}_\infty(f,Q_o) < \delta$, then f is within ε distance to a map $g : Q_o \to \ell^\infty$ that factors through a metric tree. In fact, they proved that f is withing ε distance to g such that $\mathcal{H}^{2,m}_\infty(g,Q_o) = 0$ and they used implication (e)$\Rightarrow$(a) of Theorem 106 to conclude that g factors through a metric tree.

The equivalence of conditions (b)-(f) in Theorem 106 is a consequence of a more general result:

Theorem 108. *If $f : [0,1]^{n+m} \to X$, $n \geq 1$, $m \geq 0$, is a Lipschitz map into a metric space, and $E \subset [0,1]^{n+m}$ is a measurable set, then the following conditions are equivalent:*

(b') $\operatorname{rank} \operatorname{md}(f,x) \leq n-1$ *almost everywhere in E.*

(c') $\Theta^{*n}(f,x) = 0$ *almost everywhere in E.*

(d') $\Theta^{n}_{*}(f,x) = 0$ *almost everywhere in E.*

(e') $\mathcal{H}^{n,m}_{\infty}(f,E) = 0$.

(f') $\hat{\mathcal{H}}^{n,m}_{\infty}(f,E) = 0$.

To understand the statements in Theorem 106 and Theorem 108, recall that a *metric tree*, also known as an $\mathbb{R}$-*tree*, is a geodesic space which contains no subsets homeomorphic to $\mathbb{S}^1$, so it is a geodesic space without "loops". Other equivalent definitions are explained in Section 2.5. The notion of metric derivative and its rank used in (b) and (b') was covered in Chapter 5.

The mapping densities in (c), (c') and (d), (d') were introduced in [27]. (compare to Definition 63). For a mapping $f : Q_o = [0,1]^k \to X$ into a metric space, and x in the interior of Q_o, we define the *upper* and the *lower n-densities* by

$$\Theta^{*n}(f,x) = \limsup_{r\to 0} \frac{\mathcal{H}^n_\infty(f(B(x,r)))}{\omega_n r^n}, \quad \Theta^{n}_{*}(f,x) = \liminf_{r\to 0} \frac{\mathcal{H}^n_\infty(f(B(x,r)))}{\omega_n r^n},$$

where $\mathcal{H}^n_\infty$ is the Hausdorff content and ω_n is the volume of the unit ball in $\mathbb{R}^n$.

Now, we introduce the objects in parts (e),(e') and (f), (f'). For a Lipschitz mapping $f : Q_o = [0,1]^{n+m} \to X$, $n \geq 1$, $m \geq 0$, into a metric space, Azzam and Schul [4] defined the (n,m)-*mapping content* of a set $E \subset Q_o$. However, we shall use a slightly different version of this definition that was recently introduced by David and Schul [8]:

$$\mathcal{H}^{n,m}_\infty(f,E) = \inf \sum_i \mathcal{H}^n_\infty(f(Q_i))(\operatorname{diam}(Q_i))^m,$$

where the infimum is taken over all coverings $E \subset \bigcup_i Q_i \subset Q_o$ by closed dyadic cubes Q_i. Since any two dyadic cubes either have disjoint interiors or one is a subset of another one, by removing unnecessary cubes, we may assume that all cubes in the covering have pairwise disjoint interiors.

Observe that

$$\mathcal{H}^{n,m}_{\infty}(f,E) = \mathcal{H}^{n,m}_{\infty}(f,\tilde{E}) \quad \text{if } \tilde{E} \subset E \text{ and } \mathcal{H}^{n+m}(E \setminus \tilde{E}) = 0. \tag{6.1}$$

If the coverings are allowed to be by arbitrary sets, we denote the analogous content by

$$\hat{\mathcal{H}}^{n,m}_{\infty}(f,E) = \inf \sum_i \mathcal{H}^n_{\infty}(f(A_i))(\operatorname{diam}(A_i))^m,$$

where $E \subset Q_o$, and the infimum is taken over all coverings $E \subset \bigcup_i A_i \subset Q_o$ by arbitrary sets. Obviously, for any set E,

$$\hat{\mathcal{H}}^{n,m}_{\infty}(f,E) \leq \mathcal{H}^{n,m}_{\infty}(f,E), \tag{6.2}$$

however, it is not known if the two quantities are comparable [8, Question 1.15].

We can easily see from the definitions that

$$\frac{\omega_m}{2^m}\hat{\mathcal{H}}^{n,m}_{\infty}(f,E) = \tilde{\mathcal{H}}^{n,m}_{\infty}(f,E)$$

for all f and E. But by Lemma 62,

$$\tilde{\mathcal{H}}^{n,m}_{\infty}(f,E) = \Phi^{n,m}_{\infty}(f,E)\,,$$

where the content on the right is from Definition 54.

So, we obtain the following equivalent definition for $\hat{\mathcal{H}}^{n,m}_{\infty}(f,E)$. However, we will not use this result in this chapter.

Lemma 109. *If $f : Q_o = [0,1]^{n+m} \to X$ is Lipschitz and $E \subset [0,1]^{n+m}$, then*

$$\hat{\mathcal{H}}^{n,m}_{\infty}(f,E) = \inf \sum_i \frac{\omega_n}{2^n}\big(\operatorname{diam} f(A_i)\big)^n(\operatorname{diam} A_i)^m,$$

where the infimum is taken over all coverings $E \subset \bigcup_i A_i \subset Q_o$.

In the course of the proofs we obtained other equally important results that are seemingly unrelated to the theorems listed above.

Theorem 110. *Suppose that $f : \mathbb{R}^n \supset \Omega \to X$ is a Lipschitz continuous map from an open set onto a metric space X, $f(\Omega) = X$. Then, $\dim X = n$ if and only if $\mathcal{H}^n(X) > 0$.*

Here $\dim X$ stands for the topological dimension; this is relabelled and proven as Theorem 134.

Theorem 111. *Let $\Phi : X \to Y$ be a map between metric spaces that preserves length of rectifiable curves i.e., $\ell_Y(\Phi \circ \gamma) = \ell_X(\gamma)$ for all rectifiable curves $\gamma : [a,b] \to X$. Let $f : \Omega \to X$ be a locally Lipschitz map defined on an open set $\Omega \subset \mathbb{R}^n$ for some n, and let $\tilde{X} = f(\Omega)$. Then for any Borel function $g : \tilde{X} \to [0,\infty]$ we have*

$$\int_{\tilde{X}} g(x)\, d\mathcal{H}^n(x) = \int_{\Phi(\tilde{X})} \Big(\sum_{x \in \Phi^{-1}(y) \cap \tilde{X}} g(x) \Big)\, d\mathcal{H}^n(y).$$

See Theorem 123, and Theorem 127 for a more general statement. Note that the theorem holds for any n. The result looks like the area formula under the assumption that the derivative of Φ is an *isometry*. The only problem is that under the assumptions of the theorem the derivative of Φ is not and cannot be defined.

The next result (see also Theorem 126) is a simple consequence of Theorem 111. It proves that the Hausdorff measure on a countably rectifiable subset of the Heisenberg group $\mathbb{H}^n$ equals the Lebesgue measure of projections onto $\mathbb{R}^{2n}$, taking into account the multiplicity of the projection. To our surprise, it seems that the result has not been known before.

We say that a subset $E \subset X$ of a metric space is *countably k-rectifiable* if there is a family of Lipschitz mappings $f_i : \mathbb{R}^k \supset E_i \to E$ defined on measurable sets $E_i \subset \mathbb{R}^k$ such that

$$\mathcal{H}^k\Big(E \setminus \bigcup_{i=1}^{\infty} f_i(E_i) \Big) = 0.$$

Let $\mathbb{H}^n$ be the Heisenberg group (see Section 2.6) and let $\pi : \mathbb{H}^n \to \mathbb{R}^{2n}$ be the projection onto the first $2n$ coordinates.

Theorem 112. *Assume that a set $E \subset \mathbb{H}^n$ is countably k-rectifiable for some $k \leq n$. Then for any Borel function $g : E \to [0,\infty]$, we have*

$$\int_E g(x)\, d\mathcal{H}^k_{cc}(x) = \int_{\pi(E)} \Big(\sum_{x \in \pi^{-1}(y) \cap E} g(x) \Big)\, d\mathcal{H}^k(y).$$

There are other new results included in this chapter and we follow a convention that new results are denoted as a *Theorem* or a *Proposition*, while important known results are cited as a *Lemma* or a *Corollary*.

6.1.1 Structure of the chapter

In Section 6.2 we investigate the behavior of the rank of the metric derivative under composition of maps. A new result of interest is Proposition 1.2.

In Section 6.3 we discuss mappings between metric spaces that preserve lengths of rectifiable curves and we prove Theorem 111 (Theorem 123) and Theorem 112 (corollary of Theorem 126 applied to the Heisenberg groups), as well as a more general result, Theorem 127.

In Section 6.4 we discuss applications of metric differentiabity of Lipschitz maps to topological dimension of metric spaces and we prove Theorem 110 (Theorem 134). This result is a consequence of known facts about topological dimension and the following new result (Theorem 135):

Theorem 113. *If $f : \mathbb{R}^n \supset \Omega \to X$ is a Lipschitz map from an open set to a metric space X of topological dimension $\dim X = k$, then $\operatorname{rank} \operatorname{md}(f, x) \leq k$ for almost all $x \in \Omega$.*

One of the implications in Theorem 106 is already proved in Section 6.4. Since a metric tree has topological dimension 1, it follows from Theorem 113 that a Lipschitz map f that factors through a metric tree must satisfy $\operatorname{rank} \operatorname{md}(f, x) \leq 1$ a.e. which is implication (a)⇒(b) in Theorem 106.

In Section 6.5 we discuss a well known and a general construction of a factorization of a Lipschitz map $f : X \to Y$ defined on a quasiconvex metric space. The new result is Theorem 145. This construction is used in Section 6.6 to prove implication (b)⇒(a) of Theorem 106.

Thus in Section 6.6 we use results from all previous sections and prove the following result which is the equivalence of (a) and (b) in Theorem 106.

Theorem 114. *If $f : [0, 1]^n \to X$, $n \geq 1$, is a Lipschitz map into a metric space, then f factors through a metric tree if and only if $\operatorname{rank} \operatorname{md}(f, x) \leq 1$ almost everywhere.*

Finally in Section 6.7 we prove Theorem 108 which along with Theorem 114 completes the proof of Theorem 106.

A quick note on notation: The unit ball and the unit sphere (centered at 0) in $\mathbb{R}^n$ will be denoted by $\mathbb{B}^n$ and $\mathbb{S}^{n-1}$. The (small inductive) topological dimension of X is denoted by

$\dim X$.

6.2 Rank of metric derivative – revisited

Recall the definition of the rank of the metric derivative from Section 5.2. In this section we prove several results regarding the rank of compositions of maps. Proposition 121 is a new result.

We begin by two Corollaries of the area formula, Theorem 103.

Corollary 115. *Let $f : \Omega \to X$ be a Lipschitz mapping from an open set $\Omega \subset \mathbb{R}^n$ to a metric space X. Then $\mathcal{H}^n(f(\Omega)) > 0$ if and only if* $\operatorname{rank} \operatorname{md}(f, x) = n$ *(i.e., σ is a norm) on a set of positive measure.*

Corollary 116. *If $f = (f_1, f_2, \ldots) : \mathbb{R}^n \supset \Omega \to \ell^\infty$ is Lipschitz continuous, then $\mathcal{H}^n(f(\Omega)) > 0$ if and only if there is a set $E \subset \Omega$ of positive measure and indices $i_1 < i_2 < \ldots < i_n$ such that*

$$\det \left[\frac{\partial f_{i_k}(x)}{\partial x_\ell} \right]_{1 \le k, \ell \le n} \neq 0 \quad \textit{for all } x \in E.$$

Corollary 116 follows from Corollary 115 and Lemma 98. For a direct proof of Corollary 116 that does not use Kirchheim's theorems, see [25, Theorem 2.2]. The next three lemmata will be used in the proofs of Proposition 121, Theorem 135 and Theorem 114.

Lemma 117. *Suppose that $g : \mathbb{R}^m \to \mathbb{R}^n$ is any mapping and that $f : \mathbb{R}^n \to \mathbb{R}^N$ is Lipschitz continuous. If g and $f \circ g$ are differentiable at $x \in \mathbb{R}^m$, then* $\operatorname{rank} D(f \circ g)(x) \le \operatorname{rank} Dg(x)$.

Indeed, if L is the Lipschitz constant of f, then the directional derivatives of $f \circ g$ satisfy $|D_v(f \circ g)(x)| \le L|D_v g(x)|$ and hence $\ker Dg(x) \subset \ker D(f \circ g)(x)$.

The lemma easily generalizes to the case of the metric derivative

Lemma 118. *If X, Y are metric spaces and $g : \mathbb{R}^n \supset \Omega \to X$, $f : X \to Y$ are Lipschitz mappings, then* $\operatorname{rank} \operatorname{md}(f \circ g, x) \le \operatorname{rank} \operatorname{md}(g, x)$ *for almost all $x \in \Omega$.*

Indeed, it is easily seen that $\operatorname{md}(f \circ g, x) \le L \operatorname{md}(g, x)$, whenever g and $f \circ g$ are metrically differentiable at x.

The next result is well known and it follows easily from the Brouwer fixed point theorem, see [46, Lemma 7.23].

Lemma 119. *If $h : \bar{B}^n(0,\varepsilon) \to \mathbb{R}^n$ is continuous and $|h(x) - x| < \varepsilon/2$ for all $|x| = \varepsilon$, then $\bar{B}^n(0,\varepsilon/2) \subset h(\bar{B}^n(0,\varepsilon))$.*

Remark 120. We will use Lemma 119 in the proofs of Proposition 121 and Theorem 135. The reader should compare the two proofs—finding similarities will help with a better understanding of the underlying ideas.

The next result is of independent interest and it will be used in the proof of Theorem 114.

Proposition 121. *Let $f : \Omega \to X$ be a Lipschitz mapping from an open set $\Omega \subset \mathbb{R}^n$ to a metric space X, such that* $\operatorname{rank}\operatorname{md}(f, y) \leq k$ *for almost all $y \in \Omega$. If $g : U \to \Omega$ is a Lipschitz map from an open set $U \subset \mathbb{R}^m$, then* $\operatorname{rank}\operatorname{md}(f \circ g, x) \leq k$ *for almost all $x \in U$.*

Remark 122. This result is not obvious, because it may happen that the image of g is contained in the set where f is not metrically differentiable and therefore, we cannot even try to estimate $\operatorname{rank}\operatorname{md}(f \circ g, x)$ by $\operatorname{rank}\operatorname{md}(f, g(x))$, because $\operatorname{md}(f, g(x))$ might not exist.

Proof. For simplicity assume that $U = \mathbb{R}^m$ and $\Omega = \mathbb{R}^n$. Suppose to the contrary that $\operatorname{rank}\operatorname{md}(f \circ g, \cdot) \geq k+1$ on a set of positive measure. We may assume that $X = \ell^\infty$, $f = (f_1, f_2, \ldots) : \mathbb{R}^n \to \ell^\infty$, so the rank of the componentwise derivative satisfies $\operatorname{rank} D(f \circ g) \geq k+1$ on a set of positive measure. Therefore, we may find a set $E \subset \mathbb{R}^m$ of positive measure, such that a $(k+1) \times (k+1)$ minor of $D(f \circ g)$ is non-zero in E, see Lemma 98. Without loss of generality we may assume that

$$\det \left[\frac{\partial (f_i \circ g)(x))}{\partial x_j} \right]_{1 \leq i,j \leq k+1} \neq 0 \quad \text{for all } x \in E. \tag{6.3}$$

Fix $x_o \in E$ such that g is differentiable at x_o. To complete the proof, it suffices to show that there is a neighborhood $G \subset \mathbb{R}^n$ of $g(x_o) \in \mathbb{R}^n$ such that the derivative of the mapping $F = (f_1, \ldots, f_{k+1}) : \mathbb{R}^n \to \mathbb{R}^{k+1}$ has rank $k+1$ on a set of positive measure in G, because this will imply that $\operatorname{rank}\operatorname{md}(f, \cdot) \geq k+1$, on a set of positive measure, see Lemma 98.

Without loss of generality we may assume that $x_o = 0$ and $g(x_o) = 0$. From now on we restrict g to the $(k+1)$-dimensional subspace generated by the first $(k+1)$-coordinates so we identify g with $g := g|_{\mathbb{R}^{k+1}} : \mathbb{R}^{k+1} \to \mathbb{R}^n$. Since by (6.3), $\operatorname{rank} D(F \circ g)(0) = k+1$ (because $0 = x_o \in E$), it follows from Lemma 117, that $\operatorname{rank} Dg(0) \geq k+1$, and hence $\operatorname{rank} Dg(0) = k+1$, because g is defined on $\mathbb{R}^{k+1}$. By pre-composing g with a suitable linear map and by choosing a coordinate system in $\mathbb{R}^n$ so that $Dg(0)(\mathbb{R}^{k+1})$ is the subspace of $\mathbb{R}^n$ generated by the first $(k+1)$-coordinates, we may assume that $Dg(0)$ is the identity embedding of $\mathbb{R}^{k+1}$ into $\mathbb{R}^n$. All these assumptions are made only to make notation simpler.

Thus $g : \mathbb{R}^{k+1} \to \mathbb{R}^n = \mathbb{R}^{k+1} \times \mathbb{R}^{n-k-1}$, and $g(x) = (x, 0) + o(|x|) \in \mathbb{R}^{k+1} \times \mathbb{R}^{n-k-1}$. Since by (6.3), $\det D(F \circ g)(0) \neq 0$, we may assume (after post-composing with an affine isomorphism) that $F(0) = 0$ and $D(F \circ g)(0) = \mathrm{id}$ i.e., $(F \circ g)(x) = x + o(|x|)$.

Therefore, there is $\varepsilon > 0$, such that if $|x| = \varepsilon$, then

$$|(F \circ g)(x) - x| < \frac{\varepsilon}{6} \quad \text{and} \quad |g(x) - (x,0)| < \frac{\varepsilon}{6L},$$

where L is a Lipschitz constant of F.

Fix $y \in \mathbb{R}^{n-k-1}$, $|y| < \frac{\varepsilon}{6L}$, such that F is differentiable at almost all points of the hyperplane $\mathbb{R}^{k+1} \times \{y\} \subset \mathbb{R}^n$ (by Fubini's theorem almost all $y \in \mathbb{R}^{n-k-1}$ have this property). For $|x| = \varepsilon$ we have

$$\begin{aligned} |F(x,y) - x| &\leq |F(x,y) - F(x,0)| + |F(x,0) - F(g(x))| + |(F \circ g)(x) - x| \\ &< L|y| + L|(x,0) - g(x)| + \frac{\varepsilon}{6} < \frac{\varepsilon}{2}. \end{aligned}$$

It follows from Lemma 119 that

$$\bar{B}^{k+1}(0, \varepsilon/2) \subset F(\bar{B}^{k+1}(0,\varepsilon) \times \{y\}).$$

In particular, the $(k+1)$-dimensional measure of $F(\bar{B}^{k+1}(0,\varepsilon) \times \{y\})$ is positive and it follows from the classical area formula that

$$\operatorname{rank} D\Big(F\big|_{\bar{B}^{k+1}(0,\varepsilon)\times\{y\}}\Big) = k+1$$

on a set of positive measure. Since it is true for almost all $y \in \mathbb{R}^{n-k-1}$ such that $|y| < \frac{\varepsilon}{6L}$, it follows that $\operatorname{rank} DF \geq k+1$ on a subset of $B^{k+1}(0,\varepsilon) \times B^{n-k-1}(0, \varepsilon/6L)$ of positive measure and we arrive at a contradiction. $\square$

6.3 Area formula for length preserving mappings

Let us start with a simplified and a more transparent version of the main result of this section which is Theorem 127.

Theorem 123. *Let $\Phi : X \to Y$ be a map between metric spaces that preserves length of rectifiable curves i.e., $\ell_Y(\Phi \circ \gamma) = \ell_X(\gamma)$ for all rectifiable curves $\gamma : [a,b] \to X$. Let $f : \Omega \to X$ be a locally Lipschitz map defined on an open set $\Omega \subset \mathbb{R}^n$ for some n, and let $\tilde{X} = f(\Omega)$. Then for any Borel function $g : \tilde{X} \to [0, \infty]$ we have*

$$\int_{\tilde{X}} g(x)\, d\mathcal{H}^n(x) = \int_{\Phi(\tilde{X})} \Big(\sum_{x \in \Phi^{-1}(y) \cap \tilde{X}} g(x) \Big)\, d\mathcal{H}^n(y).$$

Remark 124. We do not assume that Φ is continuous. For example, if the only rectifiable curves in X are constant ones (e.g., if X is the von Koch snowflake or a Cantor set), then *any* map $\Phi : X \to Y$ satisfies the assumptions of the theorem. However, the result is trivial in that case since $\tilde{X} = f(\Omega)$ consists of a single point (if Ω is connected). Thus the result is interesting only if there are many rectifiable curves in X.

Remark 125. Note that the theorem holds for any n. The result looks like the area formula under the assumption that the derivative of Φ is an *isometry*. The only problem is that under the assumptions of the theorem the derivative of Φ is not and cannot be defined.

Since according to Lemma 30, the projection $\pi : \mathbb{H}^n \to \mathbb{R}^{2n}$ preserves length of rectifiable curves, we obtain

Theorem 126. *Let $f : K \to \mathbb{H}^n$ be a Lipschitz map defined on a Borel set $K \subset \mathbb{R}^k$ for some $k \leq n$. Then for any Borel function $g : f(K) \to [0, \infty]$ we have*

$$\int_{f(K)} g(x)\, d\mathcal{H}^k_{cc}(x) = \int_{\pi(f(K))} \Big(\sum_{x \in \pi^{-1}(y) \cap f(K)} g(x) \Big)\, d\mathcal{H}^k(y).$$

Indeed, this follows immediately from Theorem 123 because according to Lemma 31, we may assume that f is defined on $\mathbb{R}^k$. Now, Theorem 112 follows immediately from Theorem 126, because any countably k-rectifiable subset of $\mathbb{H}^n$ is the union of countably many disjoint sets $f(K_i)$ plus a set of $\mathcal{H}^k_{cc}$-measure zero.

Theorem 123 is a straightforward consequence of the following more general result. While the statement of Theorem 127 is not as appealing as that of Theorem 123, we actually need this more general statement for the applications to results in Section 6.5, see Theorem 145.

Theorem 127. *Let $\Phi : X \to Y$ be a map between metric spaces. Assume that $\Omega \subset \mathbb{R}^n$ is open and $f : \Omega \to X$, is Lipschitz. For $x \in \Omega$ and $v \in \mathbb{S}^{n-1}$ let*

$$\gamma_{x,v}(t) = f(x+tv) : [0, d(x)] \to X, \quad \textit{where} \quad d(x) = \min\left\{\frac{1}{2}\operatorname{dist}(x, \partial\Omega), 1\right\}$$

be a family of Lipschitz curves in X. Assume that

$$\ell_Y(\Phi \circ \gamma_{x,v}|_{[0,t]}) = \ell_X(\gamma_{x,v}|_{[0,t]}) \quad \textit{for all } x \in \Omega,\ v \in \mathbb{S}^{n-1}, \textit{ and all } t \in [0, d(x)]. \tag{6.4}$$

Then

(a) $\operatorname{md}(\Phi \circ f, x) = \operatorname{md}(f, x)$ *for almost all* $x \in \Omega$.

(b) $J_n(\operatorname{md}(\Phi \circ f, x)) = J_n(\operatorname{md}(f, x))$ *for almost all* $x \in \Omega$.

Moreover, if $\tilde{X} = f(\Omega)$, then

(c) For any Borel function $g : \tilde{X} \to [0, \infty]$.

$$\int_{\tilde{X}} g(x)\, d\mathcal{H}^n(x) = \int_{\Phi(\tilde{X})} \Big(\sum_{x \in \Phi^{-1}(y) \cap \tilde{X}} g(x) \Big)\, d\mathcal{H}^n(y).$$

(d) For any Borel set $E \subset \tilde{X}$ and any Borel function $g : Y \to [0, \infty]$

$$\int_E (g \circ \Phi)(x)\, d\mathcal{H}^n(x) = \int_Y g(y) \mathcal{H}^0(\Phi^{-1}(y) \cap E)\, d\mathcal{H}^n(y).$$

Remark 128. Even if f is one-to-one and surjective, Φ need not be continuous for the claim of the theorem to be true. For example, we can have f defined on $(0,1)$ that bends the interval in a length preserving way, and glues 1 to 1/2. Since 1 is not a point in the domain $(0,1)$ the map is one-to-one and the inverse map $\Phi = f^{-1}$ is discontinuous at 1/2, but it preserves the length of curves $\gamma_{x,v}$. The lack of continuity does not create any problem in the proof, because $\Phi \circ f$ is locally Lipschitz continuous.

Proof. From (5.2) we know that $\mathrm{md}\,(f,x)(v)$ equals the speed of the curve $\gamma_{x,v}(t)$ at $t=0$. However, assumptions of the theorem provide information about length of curves so it will be convenient to express $\mathrm{md}\,(f,x)(v)$ as a derivative of the length of the curve $\gamma_{x,v}(t)$.

Given a Lipschitz curve $\gamma:[a,b]\to X$, let $s_\gamma:[a,b]\to[0,\ell(\gamma)]$, $s_\gamma(t)=\ell(\gamma|_{[a,t]})$, be the so-called *arc-length* parameter.

Lemma 129. *If $\gamma:[a,b]\to X$ is Lipschitz, then s_γ is Lipschitz and $\dot{s}_\gamma(t)=|\dot{\gamma}|(t)$ for almost all $t\in[a,b)$. That is,*

$$|\dot{\gamma}|(t)=\lim_{h\to 0^+}\frac{d(\gamma(t+h),\gamma(t))}{h}=\lim_{h\to 0^+}\frac{\ell(\gamma|_{[t,t+h]})}{h}\quad \textit{for almost all } t\in[a,b). \tag{6.5}$$

Proof. Let γ be L-Lipschitz. For $a\leq t_1\leq t_2\leq b$ we have

$$|s_\gamma(t_2)-s_\gamma(t_1)|=\ell(\gamma|_{[t_1,t_2]})\leq L|t_2-t_1|$$

which proves that s_γ is Lipschitz. In particular s_γ is differnetiale almost everywhere. Since the length of a curve connecting two points is no less than the distance between the points, $\dot{s}_\gamma\geq|\dot{\gamma}|$ almost everywhere. This and the equality

$$\int_a^b|\dot{\gamma}|(t)\,dt=\ell(\gamma)=s_\gamma(b)-s_\gamma(a)=\int_a^b\dot{s}_\gamma(t)\,dt$$

proves that $\dot{s}_\gamma=|\dot{\gamma}|$ almost everywhere which is (6.5). □

Lemma 130. *Let $\{v_i\}_{i=1}^\infty\subset\mathbb{S}^{n-1}$ be a countable and a dense subset of the (Euclidean) unit sphere. Let $f:\mathbb{R}^n\supset\Omega\to X$ be a Lipschitz mapping from an open set into a metric space. For $x\in\Omega$ and $v\in\mathbb{S}^{n-1}$ let $\gamma_{x,v}(t)=f(x+tv)$. Then for almost all $x\in\Omega$ and all $i=1,2,\ldots$ we have*

$$\mathrm{md}\,(f,x)(v_i)=\lim_{t\to 0^+}\frac{\ell(\gamma_{x,v_i}|_{[0,t]})}{t}\,.$$

Proof. For simplicity of notation assume that $\Omega = \mathbb{R}^n$. Fix $v \in \mathbb{S}^{n-1}$. It suffices to prove that

$$\mathrm{md}\,(f,x)(v) = \lim_{t\to 0^+} \frac{\ell(\gamma_{x,v}|_{[0,t]})}{t} \quad \text{for almost all } x \in \mathbb{R}^n, \tag{6.6}$$

because the result will be a straightforward consequence of the fact that the union of countably many sets of measure zero has measure zero.

Let $W = v^\perp = \{w \in \mathbb{R}^n : \langle w, v\rangle = 0\}$. Let $\Gamma : W \times \mathbb{R} \to X$ be defined by

$$\Gamma(w,t) = \Gamma_w(t) = f(w + tv).$$

According to Lemma 129, for every $w \in W$ and almost all $t \in \mathbb{R}$,

$$|\dot{\Gamma}_w|(t) = \lim_{h\to 0^+} \frac{\ell(\Gamma_w|_{[t,t+h]})}{h}. \tag{6.7}$$

The Fubini theorem implies that the set $\tilde{E} \subset W \times \mathbb{R}$ of points (w,t) for which (6.7) does not hold has measure zero. The mapping $\Phi : W \times \mathbb{R} \to \mathbb{R}^n$, $\Phi(w,t) = w + tv$ is a linear isometry and hence $E = \Phi(\tilde{E})$ has measure zero.

If $x \in \mathbb{R}^n \setminus E$, $x = \Phi(w,t) = w + tv$, then $\Gamma_w(t+h) = \gamma_{x,v}(h)$ so (6.7) yields

$$|\dot{\gamma}_{x,v}|(0) = |\dot{\Gamma}_w|(t) = \lim_{h\to 0^+} \frac{\ell(\Gamma_w|_{[t,t+h]})}{h} = \lim_{h\to 0^+} \frac{\ell(\gamma_{x,v}|_{[0,h]})}{h},$$

and (6.6) follows, because according to (5.2), for almost all $x \in \mathbb{R}^n$ we have

$$\mathrm{md}\,(f,x)(v) = \lim_{t\to 0} \frac{d(f(x+tv), f(x))}{t} = |\dot{\gamma}_{x,v}|(0).$$

The proof is complete. $\square$

Lemma 130 describes values of the seminorm $\mathrm{md}\,(f,x)$ on a countable and dense subset of $\mathbb{S}^{n-1}$ and the next lemma shows that this information completely determines a seminorm.

Lemma 131. *Let $\{v_i\}_{i=1}^\infty \subset \mathbb{S}^{n-1}$ be a countable and a dense subset. If σ_1, σ_2 are seminorms on $\mathbb{R}^n$, and $\sigma_1(v_i) = \sigma_2(v_i)$ for all $i = 1, 2, \ldots$, then $\sigma_1(v) = \sigma_2(v)$ for all $v \in \mathbb{R}^n$.*

Proof. Since $\sigma_1(tv_i) = \sigma_2(tv_i)$ for all $i \in \mathbb{N}$ and $t \in \mathbb{R}$, it follows that the equality holds on the set $E = \{tv_i : t \in \mathbb{R}, i \in \mathbb{N}\}$ that is dense in $\mathbb{R}^n$ in the Euclidan metric. It is a routine exercise to show that any seminorm σ is bounded by the Euclidean norm, that is $\sigma(v) \leq C|v|$. Therefore, $|\sigma(u) - \sigma(v)| \leq \sigma(u - v) \leq C|u - v|$. For $v \in \mathbb{R}^n$ choose $w_k \in E$ such that $|v - w_k| \to 0$. Then $\sigma(w_k) \to \sigma(v)$, so passing to the limit in $\sigma_1(w_k) = \sigma_2(w_k)$ as $k \to \infty$, yields the result. □

Now we have all we need to complete the proof of Theorem 127. Fix a countable and a dense set $\{v_i\}_{i=1}^{\infty} \subset \mathbb{S}^{n-1}$. According to Lemma 130, for almost all x and all i,

$$\mathrm{md}\,(f, x)(v_i) = \lim_{t \to 0^+} \frac{\ell_X(\gamma_{x,v_i}|_{[0,t]})}{t} \quad \text{and} \quad \mathrm{md}\,(\Phi \circ f, x)(v_i) = \lim_{t \to 0^+} \frac{\ell_Y(\Phi \circ \gamma_{x,v_i}|_{[0,t]})}{t} .$$

Therefore (6.4) implies that

$$\mathrm{md}\,(\Phi \circ f, x)(v_i) = \mathrm{md}\,(f, x)(v_i)$$

and (a) follows from Lemma 131, while (b) is an immediate consequence of (a). Finally (c) and (d) are consequences of Lemma 103. Since (d) easily follows from (c) it remains to prove (c).

Given a Borel function $g : \tilde{X} \to [0, \infty]$, let $G : \tilde{X} \to [0, \infty]$ be defined by

$$G(y) = \frac{g(y)}{\mathcal{H}^0(f^{-1}(y))} .$$

Then, we have

$$\begin{aligned}
\int_{\tilde{X}} g(y)\, d\mathcal{H}^n(y) &= \int_{\tilde{X}} G(y)\mathcal{H}^0(f^{-1}(y))\, d\mathcal{H}^n(y) \overset{(5.26)}{=} \int_{\Omega} G(f(x)) J_n(\mathrm{md}\,(f, x))\, d\mathcal{H}^n(x) \\
&= \int_{\Omega} G(f(x)) J_n(\mathrm{md}\,(\Phi \circ f, x))\, d\mathcal{H}^n(x) \overset{(5.25)}{=} \int_{\Phi(\tilde{X})} \sum_{x \in (\Phi\circ f)^{-1}(y)} G(f(x))\, d\mathcal{H}^n(y),
\end{aligned}$$

and it remains to observe that

$$(\Phi \circ f)^{-1}(y) = f^{-1}(\Phi^{-1}(y)) = \bigcup_{z \in \Phi^{-1}(y)} f^{-1}(z) = \bigcup_{z \in \Phi^{-1}(y) \cap \tilde{X}} f^{-1}(z),$$

because $f^{-1}(z) = \varnothing$ if $z \notin \tilde{X}$, and hence for $y \in \Phi(\tilde{X})$ we have

$$\begin{aligned}
\sum_{x \in (\Phi\circ f)^{-1}(y)} G(f(x)) &= \sum_{z \in \Phi^{-1}(y) \cap \tilde{X}} \sum_{x \in f^{-1}(z)} G(f(x)) \\
&= \sum_{z \in \Phi^{-1}(y) \cap \tilde{X}} G(z)\mathcal{H}^0(f^{-1}(z)) = \sum_{z \in \Phi^{-1}(y) \cap \tilde{X}} g(z).
\end{aligned}$$

The proof is complete. □

6.4 Topological dimension

The *topological dimension* (small inductive dimension) $\dim X$ of a metric space X is defined as follows:

- $\dim X$ is an integer greater than or equal to -1 or $\dim X = \infty$.
- $\dim X = -1$ if and only if $X = \varnothing$.
- $\dim X \leq n$ if every point in X has an arbitrarily small neighborhood whose boundary has dimension $\leq n-1$.
- $\dim X = n$ if $\dim X \leq n$ and it is not true that $\dim X \leq n-1$.
- $\dim X = \infty$ if $\dim X \leq n$ is false for all integers $n \geq -1$.

There are many other definitions of the topological dimension. They are equivalent to the above one if X is a separable metric space, see [16, 30].

The next result is well known; see e.g., [3, Theorem 2].

Lemma 132. *If an $\mathbb{R}$-tree T has at least two points, then* $\dim T = 1$.

Sketch of a proof. Clearly $\dim T \geq 1$, since T contains a segment. To prove that $\dim T \leq 1$ it suffices to show that boundaries of balls in T have dimension ≤ 0. It is not difficult to prove that boundaries of balls in T are ultrametric spaces and every non-empty ultrametric space has dimension 0, because balls in ultrmetric spaces are clopen. □

The following result is due to Szpilrajn [13]. For a proof see [28, Theorem 8.15]

Lemma 133. *If a metric space X satisfies $\mathcal{H}^{n+1}(X) = 0$, where $n \geq -1$ is an integer, then* $\dim X \leq n$.

On the other hand any Cantor set has topological dimension 0, but one can construct Cantor sets of infinite Hausdorff dimension. Thus in general, information about the topological dimension does not give any upper estimate for the Hausdorff dimension, except the situation described in Theorem 134.

Theorem 134. *Suppose that $f : \mathbb{R}^n \supset \Omega \to X$ is a Lipschitz continuous map from an open set onto a metric space X, $f(\Omega) = X$. Then,* $\dim X = n$ *if and only if* $\mathcal{H}^n(X) > 0$.

Proof. It follows from Lemma 133 that if $\dim X = n$, then $\mathcal{H}^n(X) > 0$. Thus it remains to show the opposite implication. Suppose that $\mathcal{H}^n(X) > 0$. Since $\mathcal{H}^{n+1}(X) = 0$ (X is a Lipschitz image of a subset of $\mathbb{R}^n$), Lemma 133 implies that $\dim X \leq n$ and it suffices to show that the inequality $\dim X \leq n-1$ is false. Suppose to the contrary that $\dim X \leq n-1$. Inequality $\mathcal{H}^n(X) > 0$ and Corollary 115 imply that $\operatorname{rank} \operatorname{md}(f,x) = n$ on a set of positive measure. Hence the result is a direct consequence of Theorem 135 applied to $\dim X = k \leq n-1$. Indeed, Theorem 135 implies that $\operatorname{rank} \operatorname{md}(f,x) \leq k \leq n-1$ a.e. which contradicts the fact that $\operatorname{rank} \operatorname{md}(f,x) = n$ on a set of positive measure. $\square$

Theorem 135. *If $f : \mathbb{R}^n \supset \Omega \to X$ is a Lipschitz map from an open set to a metric space X of topological dimension $\dim X = k$, then $\operatorname{rank} \operatorname{md}(f,x) \leq k$ for almost all $x \in \Omega$.*

Proof. The result is obvious if $k \geq n$ so we may assume that $k < n$. Let $\tilde{X} = f(\Omega)$. The space $\tilde{X}$ is separable and $\dim \tilde{X} \leq k$. Separability of $\tilde{X}$ allows us to assume that $\tilde{X} \subset \ell^\infty$ and $f = (f_1, f_2, \ldots) : \Omega \to \ell^\infty$.

Suppose to the contrary that $\operatorname{rank} \operatorname{md}(f,x) \geq k+1$ on a set of positive measure. We will arrive to a contradiction by showing that $\dim \tilde{X} > k$. To this end we shall need the following classical result [16, Theorem 1.9.3], [30, Theorem VI.4].

Lemma 136. *A separable metric space X has topological dimension less than or equal k, $k \geq 0$, if and only if for each closed set $C \subset X$ and a continuous map $h : C \to \mathbb{S}^k$, there is a continuous extension $H : X \to \mathbb{S}^k$ of h.*

Thus it remains to find a closed set $C \subset \tilde{X}$ and a continuous map $h : C \to \mathbb{S}^k$ that has no continuous extension $H : \tilde{X} \to \mathbb{S}^k$.

It follows from Lemma 98 that a certain $(k+1) \times (k+1)$ minor of the componentwise derivative Df is non-zero on a set of positive measure. After relabeling indices, translating Ω, and translating the image in ℓ^∞ we may assume that $0 \in \Omega$, $f(0) = 0$ and that the function

$$F = (f_1, \ldots, f_{k+1})\big|_{\tilde{\Omega}} : \tilde{\Omega} \to \mathbb{R}^{k+1}, \qquad \tilde{\Omega} = \Omega \cap \{(x_1, \ldots, x_{k+1}, 0, \ldots, 0)\} \subset \mathbb{R}^{k+1}$$

is differentiable at $0 \in \tilde{\Omega}$ with $\det DF(0) \neq 0$. Further, replacing f by $f \circ (DF(0))^{-1}$, we may assume that $0 \in \tilde{\Omega}$, and

$$F(0) = 0, \quad DF(0) = I \quad \text{(identity matrix)}.$$

Therefore, there is $r > 0$ such that $\bar{B}^{k+1}(0, r) \subset \tilde{\Omega}$ and

$$|F(x) - x| < r/4 \quad \text{whenever } |x| = r. \tag{6.8}$$

Let $\pi : \ell^\infty \to \mathbb{R}^{k+1}$, $\pi(x_1, x_2, \ldots) = (x_1, x_2, \ldots, x_{k+1})$ be the projection on the first $k+1$ components, so $F = \pi \circ (f|_{\tilde{\Omega}})$.

Let $C = f(S^k(0, r)) \subset \tilde{X} \subset \ell^\infty$, where $S^k(0, r) = \partial \bar{B}^{k+1}(0, r) \subset \tilde{\Omega}$. Note that if $y \in C$, then $\pi(y) \neq 0$. Indeed, $y = f(x)$, $|x| = r$ so $\pi(y) = F(x)$ and hence

$$|\pi(y)| \geq |x| - |\pi(y) - x| = r - |F(x) - x| > \frac{3r}{4} > 0.$$

Therefore,

$$h : C \to \mathbb{S}^k, \qquad h(y) = \frac{\pi(y)}{|\pi(y)|}$$

is well defined and continuous. It remains to show that there is no continuous extension $H : \tilde{X} \to \mathbb{S}^k$ of h. Suppose, by way of contradiction, that such H exists. Then

$$g : \bar{B}^{k+1}(0, 1) \to \mathbb{S}^k, \qquad g(x) = H(f(rx)) \tag{6.9}$$

is well defined and continuous.

If $|x| = 1$, then $rx \in S^k(0, r)$, so $f(rx) \in C$ and hence

$$g(x) = h(f(rx)) = \frac{F(rx)}{|F(rx)|}, \quad \text{whenever } |x| = 1.$$

It suffices now to show that the map

$$g\big|_{\mathbb{S}^k(0,1)} : \mathbb{S}^k \to \mathbb{S}^k, \qquad g(x) = \frac{F(rx)}{|F(rx)|} \tag{6.10}$$

satisfies

$$|g(x) - x| < \frac{1}{2} \quad \text{whenever } |x| = 1. \tag{6.11}$$

Indeed, since $g : \bar{B}^{k+1}(0,1) \to \mathbb{S}^k \subset \mathbb{R}^{k+1}$, Lemma 119 yields $\bar{B}^{k+1}(0,1/2) \subset g(\bar{B}^{k+1}(0,1))$ which contradicts the fact that the image of g is contained in the unit sphere.

To prove (6.11), let $|x| = 1$. It follows from (6.8) and the triangle inequality that

$$\left|1 - \frac{|F(rx)|}{r}\right| = \left||x| - \left|\frac{F(rx)}{r}\right|\right| \leq \left|\frac{F(rx)}{r} - x\right| < \frac{1}{4}.$$

Therefore,

$$|g(x) - x| \leq \left|\frac{F(rx)}{|F(rx)|} - \frac{F(rx)}{r}\right| + \left|\frac{F(rx)}{r} - x\right| < \left|\frac{F(rx)}{|F(rx)|}\Big(1 - \frac{|F(rx)|}{r}\Big)\right| + \frac{1}{4} < \frac{1}{2}.$$

This proves (6.11) and completes the proof of the theorem. □

As a corollary of Theorem 135, Lemma 132 and Corollary 115 we obtain

Theorem 137. *If $f : \mathbb{R}^n \supset \Omega \to T$, is a Lipschitz map from an open set into an $\mathbb{R}$-tree T, then* $\operatorname{rank} \operatorname{md}(f,x) \leq 1$ *a.e. If in addition, $n \geq 2$, then $\mathcal{H}^n(f(\Omega)) = 0$.*

6.5 Factoring Lipschitz maps

Material of this section is based mostly on [50]. Similar constructions appear also in [36, 44]. Theorem 145 is new.

Given metric spaces X, Y, Z, we say that a Lipschitz map $\Phi : X \to Y$ *factors through* Z, if there are Lipschitz maps $\psi : X \to Z$ and $\phi : Z \to Y$ such that $\Phi = \phi \circ \psi$.

We say that a metric space (X,d) is C_q-*quasiconvex*, where $C_q \geq 1$ is a constant, if for any $x, y \in X$ there is a rectifiable curve $\gamma : [0,1] \to X$ such that $\gamma(0) = x$, $\gamma(1) = y$ and $\ell(\gamma) \leq C_q d(x,y)$. A meric space is said to be *quasiconvex* if it is C_q-quasiconvex for some $C_q \geq 1$.

Let X be a C_q-quasiconvex metric space, Y another metric space, and let $\Phi : X \to Y$ be an L-Lipschitz map. We define a quasimetric on X by

$$d_\Phi(x,y) = \inf\{\ell(\Phi \circ \gamma) : \gamma : [0,1] \to X,\ \gamma(0) = x,\ \gamma(1) = y\}, \tag{6.12}$$

where the infimum is over all rectifiable curves γ connecting x to y in X. Note that, with a suitable reparameterization, we can assume that γ is Lipschitz.

It is easy to see that

$$d_Y(\Phi(x), \Phi(y)) \leq d_\Phi(x, y) \leq C_q L d_X(x, y). \tag{6.13}$$

In particular,

$$d_\Phi(x, y) = 0 \quad \Rightarrow \quad \Phi(x) = \Phi(y). \tag{6.14}$$

However, in general,

$$\Phi(x) = \Phi(y) \quad \not\Rightarrow \quad d_\Phi(x, y) = 0.$$

Let $\sim$ be an equivalence relation in X defined by

$$x \sim y \quad \text{if and only if} \quad d_\Phi(x, y) = 0,$$

and let $Z_\Phi = X/\sim$. We equip Z with the quotient distance

$$d_\Phi([x], [y]) := d_\Phi(x, y)$$

(one needs to check first that d_Φ is well defined in Z_Φ i.e., if $x \sim x'$ and $y \sim y'$, then $d_\Phi(x, y) = d_\Phi(x', y')$). The next result is an easy exercise left to the reader.

Lemma 138. *(Z_Φ, d_Φ) is a metric space.*

Define now the mappings

$$X \xrightarrow{\psi} Z_\Phi \xrightarrow{\phi} Y$$
$$x \xmapsto{\psi} [x] \xmapsto{\phi} \Phi(x)$$

so $\Phi = \phi \circ \psi$. Note that the mapping

$$\phi : Z_\Phi \to Y, \quad \phi([x]) = \Phi(x)$$

is well defined, because (6.14) yields

$$[x] = [x'] \quad \equiv \quad x \sim x' \quad \equiv \quad d_\Phi(x, x') = 0 \quad \Rightarrow \Phi(x) = \Phi(x').$$

Lemma 139. *The mapping $\psi : X \to Z_\Phi$ is C_qL-Lipschitz and the mapping $\phi : Z_\Phi \to Y$ is 1-Lipschitz. Therefore $\Phi : X \to Y$ factors through Z_Φ, namely $\Phi = \phi \circ \psi$.*

Proof. The mapping ψ is C_qL-Lipschitz because according to (6.13)

$$d_\Phi(\psi(x), \psi(y)) = d_\Phi([x], [y]) = d_\Phi(x, y) \leq C_qLd_X(x, y).$$

On the other hand, ϕ is 1-Lipschitz because according to (6.13)

$$d_Y(\phi([x]), \phi([y])) = d_Y(\Phi(x), \Phi(y)) \leq d_\Phi(x, y) = d_\Phi([x], [y]).$$

□

Corollary 140. *For any curve $\alpha : [0, 1] \to Z_\Phi$ we have $\ell(\phi \circ \alpha) \leq \ell(\alpha)$.*

Indeed, ϕ is 1-Lipschitz and composing with a 1-Lipschitz map cannot increase the length of a curve.

Lemma 141. *If $\gamma : [0, 1] \to X$ is a rectifiable curve and $\alpha = \psi \circ \gamma : [0, 1] \to Z_\Phi$, then $\ell(\alpha) = \ell(\phi \circ \alpha)$.*

In other words, ϕ preserves lengths of curves in Z_Φ that are images of rectifiable curves in X.

Proof. Let $\gamma : [0, 1] \to X$ and $\alpha = \psi \circ \gamma : [0, 1] \to Z_\Phi$. In view of Corollary 140 it suffices to show that

$$\ell(\phi \circ \alpha) \geq \ell(\alpha). \tag{6.15}$$

Note that

$$\phi \circ \alpha = \Phi \circ \gamma. \tag{6.16}$$

Indeed, $\phi \circ \alpha = \phi \circ \psi \circ \gamma = \Phi \circ \gamma$. Now taking the supremum over the partitions $0 = t_0 < t_1 < \ldots < t_n = 1$ we get

$$\ell(\alpha) = \sup \sum_{i=0}^{n-1} d_\Phi([\gamma(t_i)], [\gamma(t_{i+1})]) = \sup \sum_{i=0}^{n-1} d_\Phi(\gamma(t_i), \gamma(t_{i+1})) \leq \heartsuit.$$

Since $\gamma|_{[t_i,t_{i+1}]}$ is a rectifiable curve connecting $\gamma(t_i)$ to $\gamma(t_{i+1})$, the definition of d_Φ (see (6.12)) yields that

$$d_\Phi(\gamma(t_i),\gamma(t_{i+1})) \leq \ell\left((\Phi\circ\gamma)|_{[t_i,t_{i+1}]}\right)$$

and hence

$$\heartsuit \leq \sup\sum_{i=0}^{n-1}\ell\left((\Phi\circ\gamma)|_{[t_i,t_{i+1}]}\right) = \ell(\Phi\circ\gamma) \stackrel{(6.16)}{=} \ell(\phi\circ\alpha),$$

This completes the proof of (6.15) and hence that of the lemma. □

Corollary 142. *(Z_Φ, d_Φ) is a length space. If in addition, X is compact, then Z_Φ is a geodesic space.*

Proof. If $[x],[y]\in Z_\Phi$ and $\gamma:[0,1]\to X$, $\gamma(0)=x$, $\gamma(1)=y$, is a rectifiable curve, then $\alpha=\psi\circ\gamma:[0,1]\to Z_\Phi$, $\alpha(0)=[x]$, $\alpha(1)=[y]$ and according to Lemma 141 and (6.16),

$$\ell(\alpha)=\ell(\phi\circ\alpha)=\ell(\Phi\circ\gamma).$$

Therefore, the definition of d_Φ yields

$$d_\Phi([x],[y]) = d_\Phi(x,y) = \inf_\gamma \ell(\Phi\circ\gamma) = \inf_\alpha \ell(\alpha). \tag{6.17}$$

We proved that $d_\Phi([x],[y])$ equals the infimum of length of curves α in Z_Φ that have a special form $\alpha=\psi\circ\gamma$. But the infimum over all curves in Z_Φ that connect $[x]$ to $[y]$ cannot be smaller than $d_\Phi([x],[y])$ so $d_\Phi([x],[y])$ is equal to the infimum over all curves in Z_Φ that connect $[x]$ to $[y]$.

Now suppose that additionally X is compact. Then Z_Φ is also compact as a C_qL-Lipschitz image of X and hence it is geodesic by Corollary 27. □

In (6.17) we proved that the distance in Z_Φ is obtained as infimum of lengths over a subclass of curves $\alpha=\psi\circ\gamma$ connecting given two points. While, in general, not every rectifiable curve in Z_Φ is of that form (see Example 144), all rectifiable curves in Z_Φ can be well approximated by such curves.

Lemma 143. *Let $\alpha:[0,1]\to Z_\Phi$ be a Lipschitz curve. Then, there is a sequence of Lipschitz curves $\gamma_n:[0,1]\to X$, such that $\alpha_n=\psi\circ\gamma_n:[0,1]\to Z_\Phi$ satisfies: $\alpha_n(0)=\alpha(0)$, $\alpha_n(1)=\alpha(1)$, $\alpha_n\to\alpha$ uniformly, and $\ell(\alpha_n)\to\ell(\alpha)$.*

Example 144. We shall construct an example in which we necessarily have $\ell(\gamma_n) \to \infty$. In such a case, it is not possible to construct a Lipschitz curve $\gamma : [0,1] \to X$ satisfying $\alpha = \psi \circ \gamma$.

Let $X = [0,3]$, $Y = [0,2]$, and let $\Phi : X \to Y$ be defined by

$$\Phi(x) = \begin{cases} x & \text{if } 0 \le x \le 1 \\ 1 & \text{if } 1 \le x \le 2 \\ x-1 & \text{if } 2 \le x \le 3. \end{cases}$$

Then $Z_\Phi = Y$, $\psi = \Phi$, and $\phi = \mathrm{id}$. If $\alpha : [0,1] \to [0,2] = Z_\Phi$ is a Lipschitz curve such that for some $s < t$, $\alpha(s) < 1$, $\alpha(t) > 1$, then for all sufficiently large n, $\gamma_n(s) < 1$, $\gamma_n(t) > 2$, and hence $\ell(\gamma_n|_{[s,t]}) > 1$. Therefore, if α is a highly oscillating curve that crosses the point $1 \in [0,2] = Z_\Phi$ infinitely many times, we necessarily have $\ell(\gamma_n) \to \infty$.

Proof of Lemma 143. Choose a sequence of partitions $0 = t_{n,0} < t_{n,1} < \ldots < t_{n,k_n} = 1$, such that

$$\ell(\alpha) - \frac{1}{n} \le \sum_{i=0}^{k_n-1} d_\Phi(\alpha(t_{n,i}), \alpha(t_{n,i+1})) \le \ell(\alpha), \quad \Delta_n = \max_i |t_{n,i+1} - t_{n,i}| \overset{n\to\infty}{\longrightarrow} 0. \tag{6.18}$$

Fix $x_{n,i} \in X$ satisfying $[x_{n,i}] = \psi(x_{n,i}) = \alpha(t_{n,i})$. Since

$$d_\Phi(\alpha(t_{n,i}), \alpha(t_{n,i+1})) = d_\Phi(x_{n,i}, x_{n,i+1}),$$

the definition of d_Φ yields the existence of a curve

$$\gamma_{n,i} : [t_{n,i}, t_{n,i+1}] \to X, \quad \gamma_{n,i}(t_{n,i}) = x_{n,i}, \quad \gamma_{n,i}(t_{n,i+1}) = x_{n,i+1}$$

such that

$$d_\Phi(\alpha(t_{n,i}), \alpha(t_{n,i+1})) \le \ell(\Phi \circ \gamma_{n,i}) \le d_\Phi(\alpha(t_{n,i}), \alpha(t_{n,i+1})) + \frac{1}{nk_n}. \tag{6.19}$$

According to Lemma 141,

$$\ell(\Phi \circ \gamma_{n,i}) = \ell(\phi \circ (\psi \circ \gamma_{n,i})) = \ell(\psi \circ \gamma_{n,i}). \tag{6.20}$$

Therefore, (6.18), (6.19), and (6.20) yield

$$\ell(\alpha) - \frac{1}{n} \le \sum_{i=0}^{k_n-1} \ell(\psi \circ \gamma_{n,i}) \le \ell(\alpha) + \frac{1}{n}. \tag{6.21}$$

For each n, define a Lipschitz curve $\gamma_n : [0,1] \to X$ as the concatenation of the curves $\{\gamma_{n,i}\}_{i=0}^{k_n-1}$, and let $\alpha_n = \psi \circ \gamma_n$. Then (6.21) implies that $\ell(\alpha) - 1/n \le \ell(\alpha_n) \le \ell(\alpha) + 1/n$ which proves that $\ell(\alpha_n) \to \ell(\alpha)$.

Note that

$$\alpha_n(t_{n,i}) = (\psi \circ \gamma_{n,i})(t_{n,i}) = \psi(x_{n,i}) = \alpha(t_{n,i}), \tag{6.22}$$

and in particular $\alpha_n(0) = \alpha(0)$, $\alpha_n(1) = \alpha(1)$. It remains to prove that $\alpha_n \to \alpha$ uniformly.

Assume that α is M-Lipschitz. Note that $\alpha_n|_{[t_{n,i},t_{n,i+1}]} = \psi \circ \gamma_{n,i}$, so (6.20) and (6.19) yield

$$\ell(\alpha_n|_{[t_{n,i},t_{n,i+1}]}) \le d_\Phi(\alpha(t_{n,i}), \alpha(t_{n,i+1})) + \frac{1}{nk_n} \le M|t_{n,i+1} - t_{n,i}| + \frac{1}{nk_n}\,.$$

Since by (6.22), curves α_n and α coincide at the endpoints of the interval $[t_{n,i}, t_{n,i+1}]$, for $t_{n,i} \le t \le t_{n,i+1}$ we have

$$\begin{aligned} &d_\Phi(\alpha_n(t), \alpha(t)) \le \ell(\alpha|_{[t_{n,i},t_{n,i+1}]}) + \ell(\alpha_n|_{[t_{n,i},t_{n,i+1}]}) \\ &\le M|t_{n,i+1} - t_{n,i}| + \Big(M|t_{n,i+1} - t_{n,i}| + \frac{1}{nk_n}\Big) \le 2M\Delta_n + \frac{1}{nk_n} \xrightarrow{n\to\infty} 0. \end{aligned}$$

This proves uniform convergence $\alpha_n \to \alpha$ and completes the proof. □

The next result shows that for any n, the mapping $\Phi : Z_\Phi \to Y$ preserves the $\mathcal{H}^n$ measure of certain subsets of Z_Φ.

Theorem 145. *Let $\Phi : X \to Y$ be a Lipschitz map between a quasiconvex metric space X, and another metric space Y. Let $\psi : X \to Z_\Phi$ and $\phi : Z_\Phi \to Y$ be as above.*

If $f : \Omega \to X$ is a Lipschitz map defined on an open set $\Omega \subset \mathbb{R}^n$ for some n, and $\tilde{X} = f(\Omega)$, then

(a) $\operatorname{md}(\psi \circ f, x) = \operatorname{md}(\Phi \circ f, x)$ *for almost all $x \in \Omega$.*

(b) For any Borel function $g : \psi(\tilde{X}) \to [0,\infty]$,

$$\int_{\psi(\tilde{X})} g(x)\, d\mathcal{H}^n(x) = \int_{\Phi(\tilde{X})} \Big(\sum_{x \in \phi^{-1}(y) \cap \psi(\tilde{X})} g(x)\Big)\, d\mathcal{H}^n(y).$$

Proof. The result follows immediately from Theorem 127, because according to Lemma 141, the mapping $\phi : Z_\Phi \to Y$ preserves length of curves $\gamma_{x,v}(t) = (\psi \circ f)(x + tv)$. □

6.6 Proof of Theorem 114

The implication ($\Rightarrow$) easily follows from Theorem 137 and from Lemma 118. Therefore, it remains to prove the implication ($\Leftarrow$). Some ideas used below are based on the proof of Theorem 5 in [50].

Since the proof will rely on the characterization of $\mathbb{R}$-trees as in part (g) of Lemma 29, we need a few simple lammata about the integral expression there, but first we will start with an informal heuristic discussion.

If $D \subset \mathbb{R}^2$ is a smooth bounded simply connected domain, and $\gamma : \mathbb{S}^1 \to \partial D$ is an orientation preserving parametrization of the boundary, then the area of D can be expressed as

$$A(D) = \int_D dx \wedge dy = \int_{\partial D} x\, dy = \int_{\mathbb{S}^1} \gamma^*(x\, dy).$$

If $\gamma : \mathbb{S}^1 \to \mathbb{R}^2$ is any Lipschitz curve, then

$$A(\gamma) := \int_{\mathbb{S}^1} \gamma^*(x\, dy)$$

represents the *oriented area enclosed by* γ—the sum (possibly infinite) of areas of bounded connected components of $\mathbb{R}^2 \setminus \gamma(\mathbb{S}^1)$, multipled by the corresponding winding numbers.

Thus, roughly speaking, condition (g) in Lemma 29 says that there are no closed curves in X with non-trivial "holes", as otherwise we could project such a curve to $\mathbb{R}^2$, by composing it with a suitably constructed Lipschitz map $\pi : X \to \mathbb{R}^2$, so that the resulting curve $\pi \circ \gamma$ would bound a non-zero oriented area

$$A(\pi \circ \gamma) = \int_{\mathbb{S}^1} (\pi \circ \gamma)^*(x\, dy) \neq 0.$$

This interpretation is consistent with our intuition that $\mathbb{R}$-trees are geodesic spaces without non-trivial loops.

In fact, we will not use the above geometric interpretation of $A(\gamma)$, but it is important to keep it in mind to provide intuition for the estimates that we do next.

Lemma 146. *Let $\gamma\colon \mathbb{S}^1 \to \mathbb{R}^2$ be a Lipschitz curve and let $\Gamma = (\Gamma_1, \Gamma_2)\colon \bar{\mathbb{B}}^2 \to \mathbb{R}^2$ be any Lipschitz extension of γ to the closed unit disk. Then*

$$A(\gamma) = \int_{\mathbb{B}^2} d\Gamma_1 \wedge d\Gamma_2 = \int_{\mathbb{B}^2} \det D\Gamma. \tag{6.23}$$

If Γ is smooth up to the boundary, it follows from Stokes' theorem, and in the Lipschitz case one can prove it by using a smooth approximation. For a detailed proof of a more general result, see for example, [11, Lemma 4.9].

If $\gamma : \mathbb{S}^1 \to \mathbb{R}^2$ is a Lipschitz curve and $\bar{\gamma} = (\bar{\gamma}_1, \bar{\gamma}_2) : [0,1] \to \mathbb{R}^2$, $\bar{\gamma}(t) = \gamma(\exp(2\pi i t))$, then $\bar{\gamma}(0) = \bar{\gamma}(1)$, and

$$A(\gamma) = \int_{\mathbb{S}^1} \gamma^*(x\, dy) = \int_0^1 \bar{\gamma}_1(t)\bar{\gamma}_2'(t)\, dt,$$

so with a slight abuse of notation we can identify γ with $\bar{\gamma}$ and, whenever it is convenient, regard γ as a curve $\gamma : [0,1] \to \mathbb{R}^2$ satisfying $\gamma(0) = \gamma(1)$.

Lemma 147. *Suppose that $\gamma_n\colon \mathbb{S}^1 \to \mathbb{R}^2$, $n = 1, 2, \dots$ are Lipschitz curves that uniformly converge to a Lipschitz curve $\gamma\colon \mathbb{S}^1 \to \mathbb{R}^2$. Assume also that there exists $M > 0$ such that $\ell(\gamma_n) \leq M$ for all n. Then,*

$$\lim_{n\to\infty} A(\gamma_n) = A(\gamma)\,.$$

Proof. If we write $\gamma_n = (\gamma_{n,1}, \gamma_{n,2})$, then

$$\begin{aligned} A(\gamma) - A(\gamma_n) &= \int_0^1 \gamma_1\gamma_2'\, dt - \int_0^1 \gamma_{n,1}\gamma_{n,2}'\, dt \\ &= \int_0^1 (\gamma_1 - \gamma_{n,1})\gamma_2'\, dt + \int_0^1 \gamma_{n,1}(\gamma_2' - \gamma_{n,2}')\, dt \\ &= \int_0^1 (\gamma_1 - \gamma_{n,1})\gamma_2'\, dt - \int_0^1 \gamma_{n,1}'(\gamma_2 - \gamma_{n,2})\, dt\,, \end{aligned}$$

where the last equality follows from the integration by parts. If we show that the the right-hand side converges to zero we are done. Since $|\gamma'|$ is bounded and $\gamma_{n,1} \to \gamma_1$ uniformly, it follows that

$$\left| \int_0^1 (\gamma_1 - \gamma_{n,1})\gamma_2'\, dt \right| \leq \|\gamma'\|_\infty \int_0^1 |\gamma_1 - \gamma_{n,1}|\, dt \overset{n\to\infty}{\longrightarrow} 0.$$

Next, we have

$$\left| \int_0^1 \gamma_{n,1}'(\gamma_2 - \gamma_{n,2})\, dt \right| \leq \|\gamma_2 - \gamma_{n,2}\|_\infty \int_0^1 |\gamma_{n,1}'|\, dt \leq M\|\gamma_2 - \gamma_{n,2}\|_\infty \overset{n\to\infty}{\longrightarrow} 0.$$

The proof is complete. $\square$

For the rest of this section, $f\colon [0,1]^n \to X$ is Lipschitz and $Z := Z_f$ is the corresponding metric space constructed as in Section 6.5, together with Lipschitz maps $\psi\colon [0,1]^n \to Z$ and $\phi\colon Z \to [0,1]^n$, such that $f = \phi \circ \psi$.

Now, we are ready to prove the implication ($\Leftarrow$). We need to prove that if $f : [0,1]^n \to X$ is Lipschitz, and $\operatorname{rank} \operatorname{md}(f, x) \leq 1$ a.e., then f factors through an $\mathbb{R}$-tree. Since it factors through Z, it suffices to prove that Z is an $\mathbb{R}$-tree.

By Corollary 142, Z is geodesic, so according to part (g) of Lemma 29, it suffices to show that for every Lipschitz curve $\alpha\colon \mathbb{S}^1 \to Z$ and every Lipschitz function $\pi\colon Z \to \mathbb{R}^2$,

$$A(\pi \circ \alpha) = \int_{\mathbb{S}^1} (\pi \circ \alpha)^*(x\, dy) = 0. \tag{6.24}$$

First, assume that $\alpha = \psi \circ \gamma : \mathbb{S}^1 \to Z$, where $\gamma : \mathbb{S}^1 \to [0,1]^n$ is a Lipschitz curve. Let $g : \bar{\mathbb{B}}^2 \to [0,1]^n$ be a Lipschitz extension of γ to the closed unit disc.

For technical reasons that will be explained later, we need an extension g with the property that it maps the interior of the disc to the interior of the cube $g(\mathbb{B}^2) \subset (0,1)^n$. That however, can be easily guaranteed. Indeed, if $g : \bar{\mathbb{B}}^2 \to [0,1]^n$ is any Lipschitz extension of α, then $\tilde{g} : \bar{\mathbb{B}}^2 \to [0,1]^n$ defined by

$$\tilde{g}(x) = \big(g(x) - (1/2, 1/2, \ldots, 1/2)\big)|x| + (1/2, 1/2, \ldots, 1/2),$$

agrees with g on the boundary of the disc, $|x| = 1$ (and hence $\tilde{g}$ is an extension of α), and when $|x| < 1$, $\tilde{g}(x)$ is in the interior of the segment connecting the center of the cube $(1/2, 1/2, \ldots, 1/2)$ to $g(x)$ so $\tilde{g}(x)$ belongs to the interior $(0,1)^n$ of the cube.

Then, $\pi \circ \psi \circ g : \bar{\mathbb{B}}^2 \to \mathbb{R}^2$ is a Lipschitz extension of $\pi \circ \psi \circ \gamma : \mathbb{S}^1 \to \mathbb{R}^2$, and Lemma 146 yields

$$A(\pi \circ \alpha) = A(\pi \circ \psi \circ \gamma) = \int_{\mathbb{B}^2} \det D(\pi \circ \psi \circ g)\,. \tag{6.25}$$

Clearly, for Lipschitz mappings $h : \mathbb{R}^n \supset \Omega \to \mathbb{R}^m$, $\operatorname{rank} Dh(x) = \operatorname{rank} \operatorname{md}(h, x)$, whenever h is differentiable at x, so Lemma 118, and part (a) of Theorem 145 give

$$\operatorname{rank} D(\pi \circ \psi \circ g) = \operatorname{rank} \operatorname{md}(\pi \circ \psi \circ g) \leq \operatorname{rank} \operatorname{md}(\psi \circ g) = \operatorname{rank} \operatorname{md}(f \circ g) \text{ a.e.} \tag{6.26}$$

Since by assumptions, $\operatorname{rank}\operatorname{md}(f) \leq 1$ a.e., Proposition 121 implies that $\operatorname{rank}\operatorname{md}(f \circ g) \leq 1$ a.e. (this is where we use the assumption that $g(\mathbb{B}^2) \subset (0,1)^n$), and hence $\operatorname{rank} D(\pi\circ\psi\circ g) \leq 1$ by (6.26). This and (6.25) proves that $A(\pi \circ \alpha) = 0$.

That is, we proved (6.24) for curves $\alpha : \mathbb{S}^1 \to Z$ that factor through $[0,1]^n$, $\alpha = \psi \circ \gamma$, while we need to prove (6.24) for all Lipschitz curves $\alpha : \mathbb{S}^1 \to Z$. We can however, easily pass to the general case with the help of Lemmata 143 and 147.

Let $\alpha : \mathbb{S}^1 \to Z$ be a Lipschitz curve. Let $\alpha_j = \psi \circ \gamma_j$ be the approximation described in Lemma 143. Note that the curves α_j are closed because they coincide with α at the endpoints.

Then $A(\pi \circ \alpha_j) = 0$, because of the special form of α_j. Since $\alpha_j \to \alpha$ uniformly, and curves α_j have uniformly bounded length, curves $\pi \circ \alpha_j : \mathbb{S}^1 \to \mathbb{R}^2$ have uniformly bounded length and they converge uniformly to $\pi \circ \alpha$, so Lemma 147 yields

$$0 = A(\pi \circ \alpha_j) \to A(\pi \circ \alpha)$$

and (6.24) follows. The proof is complete.

6.7 Proof of Theorem 108

We will prove the theorem by proving implications $(b') \Rightarrow (c') \Rightarrow (d') \Rightarrow (e') \Rightarrow (f') \Rightarrow (b')$.

Implication $(b') \Rightarrow (c')$. According to Proposition 99, almost all $x \in E \cap (0,1)^{n+m}$ have the property that for any $j \in \mathbb{N}$, and for all sufficiently small $r > 0$, the set $f(B(x,r))$ can be covered by j^k balls each of radius $3\sqrt{k}Lr/j$, where $L = \operatorname{Lip}(f)$ and $k = \operatorname{rank}\operatorname{md}(f,x) \leq n-1$. Since $3\sqrt{k}Lr/j \leq 3\sqrt{n}Lr/j$, for all sufficiently small $r > 0$, we have

$$\mathcal{H}^n_\infty(f(B(x,r))) \lesssim_{n,L} j^k \left(\frac{r}{j}\right)^n, \quad \text{so} \quad \Theta^{*n}(f,x) \lesssim_{n,L} \frac{1}{j^{n-k}}\,.$$

Since $n - k \geq n - (n-1) = 1$, and j was arbitrary, we have $\Theta^{*n}(f,x) = 0$. Since, this is true for almost all $x \in E \cap (0,1)^{n+m}$, (c') is proved.

Implication $(c') \Rightarrow (d')$ is obvious.

Implication $(d') \Rightarrow (e')$ is a direct consequence of the following estimate.

Proposition 148. *If $f : Q_o = [0,1]^{n+m} \to X$, $n \geq 1$, $m \geq 0$, is a Lipschitz map into a metric space and $E \subset Q_o$ is a measurable set, then*

$$\mathcal{H}^{n,m}_\infty(f,E) \lesssim_{n,m} \int_E \Theta^n_*(f,x)\, d\mathcal{H}^{n+m}(x)\,.$$

Remark 149. This is a slight improvement of [27, Proposition 5.1]. The proof presented below is similar to the one in [27], but Proposition 5.1 in [27] involved a slightly different definition of $\mathcal{H}^{n,m}_\infty$ than the one we use now, and this is one of the reasons for providing details.

Proof. The function $\Theta^n_*(f,\cdot)$ is integrable as bounded and measurable. Let

$$A = \{x \in E \cap (0,1)^{n+m} : x \text{ is a Lebesgue point of } \Theta^n_*(f,\cdot)\}.$$

Fix $\varepsilon > 0$. There is an open set $U \subset Q_o$, such that $A \subset U$ and

$$\int_U \Theta^n_*(f,x)\, d\mathcal{H}^{n+m}(x) < \int_A \Theta^n_*(f,x)\, d\mathcal{H}^{n+m}(x) + \varepsilon = \int_E \Theta^n_*(f,x)\, d\mathcal{H}^{n+m}(x) + \varepsilon.$$

Let $x \in A$. By the definition of $\Theta^n_*(f,x)$, there is a sequence $r^i_x \searrow 0$ such that

$$\frac{\mathcal{H}^n_\infty(f(B(x,r^i_x)))}{\omega_n (r^i_x)^n} < \Theta^n_*(f,x) + \varepsilon, \quad B(x,r^i_x) \subset U.$$

For each i, we can find a closed dyadic cube Q^i_x such that

$$x \in Q^i_x \subset B(x,r^i_x) \quad \text{and} \quad r^i_x \lesssim_{n,m} \operatorname{diam} Q^i_x,$$

so

$$\frac{\mathcal{H}^n_\infty(f(Q^i_x))}{\omega_n (r^i_x)^n} < \Theta^n_*(f,x) + \varepsilon, \quad Q^i_x \subset U. \tag{6.27}$$

Since averages of $\Theta^n_*(f,\cdot)$ over the cubes Q^i_x converge to $\Theta^n_*(f,x)$, as $i \to \infty$, by assuming that all r^i_x are sufficiently small, we may guarantee that

$$\Theta^n_*(f,x) < \fint_{Q^i_x} \Theta^n_*(f,y)\, d\mathcal{H}^{n+m}(y) + \varepsilon. \tag{6.28}$$

Hence (6.27) and (6.28) yield

$$\mathcal{H}^n_\infty(f(Q^i_x))(\operatorname{diam} Q^i_x)^m \lesssim_{n,m} \int_{Q^i_x} \Theta^n_*(f,y)\, d\mathcal{H}^{n+m}(y) + 2\varepsilon(\operatorname{diam} Q^i_x)^{n+m}.$$

The collection of (closed) dyadic cubes

$$\mathcal{Q} = \{Q_x^i : x \in A,\ i \in \mathbb{N}\}$$

forms a covering of A. Dyadic cubes have an amazing property that given two dyadic cubes, they have disjoint interiors or one is contained in another one. Thus leaving in $\mathcal{Q}$ only the largest cubes that are not contained in any larger cube from $\mathcal{Q}$, we obtain a subfamily $\{Q_j\}_j \subset \mathcal{Q}$ of dyadic cubes with pairwise disjoint interiors such that $A \subset \bigcup_j Q_j$. Since $\mathcal{H}^{n+m}(E \setminus A) = 0$, the definition of $\mathcal{H}^{n,m}_\infty$ along with (6.1) yield

$$\begin{aligned}
&\mathcal{H}^{n,m}_\infty(f,E) = \mathcal{H}^{n,m}_\infty(f,A) \le \sum_j \mathcal{H}^n_\infty(f(Q_j))(\operatorname{diam} Q_j)^m \\
&\lesssim_{n,m} \sum_j \int_{Q_j} \Theta^n_*(f,y)\, d\mathcal{H}^{n+m}(y) + 2\varepsilon \sum_j (\operatorname{diam} Q_j)^{n+m} \\
&\lesssim_{n,m} \int_U \Theta^n_*(f,y)\, d\mathcal{H}^{n+m}(y) + 2\varepsilon \le \int_E \Theta^n_*(f,y)\, d\mathcal{H}^{n+m}(y) + 3\varepsilon
\end{aligned}$$

and the result follows upon letting $\varepsilon \to 0$. $\square$

Implication $(e') \Rightarrow (f')$ is obvious because of (6.2).

Implication $(f') \Rightarrow (b')$. According to Lemmata 37 and 98 it suffices to prove the following result.

Proposition 150. *If $f = (f_1, f_2, \ldots)\colon Q_o = [0,1]^{n+m} \to \ell^\infty$, $n \ge 1$, $m \ge 0$, is Lipschitz, $E \subset Q_o$ is measurable, and $\hat{\mathcal{H}}^{n,m}_\infty(f,E) = 0$, then* $\operatorname{rank} Df(x) \le n-1$, *for $\mathcal{H}^{n+m}$-almost all $x \in E$.*

Proof. We shall ignore the points on the boundary of the cube. If $\operatorname{rank} Df(x) \ge n$ at some x, then there exist indices $i_1 < i_2 < \ldots < i_n$ such that $\nabla f_{i_1}(x), \ldots, \nabla f_{i_n}(x)$ are linearly independent. If we let $\pi : \ell^\infty \to \mathbb{R}^n$ be the projection $\pi(x) = (x_{i_1}, \ldots, x_{i_n})$, then the latter statement is equivalent to $\operatorname{rank} D(\pi \circ f)(x) = n$. There are countably many possibilities for choosing n natural numbers, so, the proposition will follow once we prove that for any $i_1 < i_2 < \ldots < i_n$, the equality $\operatorname{rank} D(\pi \circ f)(x) = n$ occurs only on a null subset of E, i.e. $\operatorname{rank} D(\pi \circ f)(x) \le n-1$, for $\mathcal{H}^{n+m}$-almost all $x \in E$.

But this follows immediately from the next lemma since by easy estimates[1]

$$\hat{\mathcal{H}}^{n,m}_\infty(\pi\circ f, E) \le (\operatorname{Lip}\pi)^n \hat{\mathcal{H}}^{n,m}_\infty(f,E) = 0.$$

We need to apply the next lemma to $F = \pi\circ f : Q_o \to \mathbb{R}^n$.

Lemma 151. *If $F : Q_o = [0,1]^{n+m} \to \mathbb{R}^n$, $n \ge 1$, $m \ge 0$ is Lipschitz, $E \subset Q_o$ is measurable, and $\hat{\mathcal{H}}^{n,m}_\infty(F,E) = 0$, then* $\operatorname{rank} DF(x) \le n-1$ *for $\mathcal{H}^{n+m}$-almost all $x \in E$.*

Proof. According to the classical co-area formula (Lemma 104)

$$\int_E |J_F(x)|\, d\mathcal{H}^{n+m}(x) = \int_{\mathbb{R}^n} \mathcal{H}^m(F^{-1}(y)\cap E)\, d\mathcal{H}^n(y).$$

Note that if F is differentiable at x, then $|J_F(x)| = 0$ if and only if $\operatorname{rank} DF(x) \le n-1$. Therefore, it suffices to show that $|J_F(x)| = 0$ for almost all $x \in E$. To this end, it suffices to show that $\mathcal{H}^m(F^{-1}(y)\cap E) = 0$ for $\mathcal{H}^n$-almost all $y \in \mathbb{R}^n$.

Assume that $m \ge 1$. A similar argument works for $m = 0$.

Let $\varepsilon > 0$ be given. Since $\hat{\mathcal{H}}^{n,m}_\infty(F,E) = 0$, it follows that there is a covering $E \subset \bigcup_i A_i \subset Q_o$, such that

$$\sum_i \mathcal{H}^n_\infty(F(A_i))(\operatorname{diam} A_i)^m < \varepsilon.$$

According to Lemmata 20 and 18, there are Borel sets $Z_i \subset \mathbb{R}^n$ such that $F(A_i) \subset Z_i$ and $\mathcal{H}^n_\infty(F(A_i)) = \mathcal{H}^n(F(A_i)) = \mathcal{H}^n(Z_i)$. We have

$$\begin{aligned}\mathcal{H}^m_\infty(F^{-1}(y)\cap E) &\lesssim_m \sum_i (\operatorname{diam}(F^{-1}(y)\cap A_i))^m = \sum_i (\operatorname{diam}(F^{-1}(y)\cap A_i))^m \chi_{F(A_i)}(y)\\ &\le \sum_i (\operatorname{diam} A_i)^m \chi_{Z_i}(y).\end{aligned}$$

Since the function on the right hand side is Borel, integration yields

$$\int^*_{\mathbb{R}^n} \mathcal{H}^m_\infty(F^{-1}(y)\cap E)\, d\mathcal{H}^n(y) \lesssim_m \sum_i (\operatorname{diam} A_i)^m \mathcal{H}^n(Z_i) = \sum_i (\operatorname{diam} A_i)^m \mathcal{H}^n_\infty(F(A_i)) < \varepsilon.$$

Since this is true for any $\varepsilon > 0$, we conclude that

$$\int^*_{\mathbb{R}^n} \mathcal{H}^m_\infty(F^{-1}(y)\cap E)\, d\mathcal{H}^n(y) = 0.$$

Thus, $\mathcal{H}^m_\infty(F^{-1}(y)\cap E) = 0$, and hence $\mathcal{H}^m(F^{-1}(y)\cap E) = 0$, for $\mathcal{H}^n$-almost all $y \in \mathbb{R}^n$. The proof of Lemma 151 is complete. $\square$

[1]Note that π is a Lipschitz map with the Lipschitz constant $\sqrt{n}$.

This also completes the proof of Proposition 150 and hence that proof of the implication.

$\square$

This was the last implication to prove and therefore, the proof of Theorem 108 is complete. $\square$

Bibliography

[1] Luigi Ambrosio and Bernd Kirchheim. Currents in metric spaces. *Acta Math.*, 185(1):1–80, 2000.

[2] Luigi Ambrosio and Paolo Tilli. *Topics on analysis in metric spaces*, volume 25 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2004.

[3] P. D. Andreev and V. N. Berestovskiĭ. Dimensions of $\mathbb{R}$-trees and self-similar fractal spaces of nonpositive curvature [translation of mr2301597]. *Siberian Adv. Math.*, 17(2):79–90, 2007.

[4] Jonas Azzam and Raanan Schul. Hard Sard: quantitative implicit function and extension theorems for Lipschitz maps. *Geom. Funct. Anal.*, 22(5):1062–1123, 2012.

[5] Yu. D. Burago and V. A. Zalgaller. *Geometric inequalities*, volume 285 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1988. Translated from the Russian by A. B. Sosinskiĭ, Springer Series in Soviet Mathematics.

[6] Ian Chiswell. *Introduction to Λ-trees*. World Scientific Publishing Co., Inc., River Edge, NJ, 2001.

[7] Bernard Dacorogna. *Direct methods in the calculus of variations*, volume 78 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2008.

[8] Guy C David and Raanan Schul. Quantitative decompositions of lipschitz mappings into metric spaces. *arXiv preprint arXiv:2002.10318*, 2020.

[9] Guy C. David and Raanan Schul. Lower bounds on mapping content and quantitative factorization through trees, 2021. Preprint.

[10] Roy O. Davies. Increasing sequences of sets and Hausdorff measure. *Proc. London Math. Soc. (3)*, 20:222–236, 1970.

[11] Noel Dejarnette, Piotr Hajł asz, Anton Lukyanenko, and Jeremy T. Tyson. On the lack of density of Lipschitz mappings in Sobolev spaces with Heisenberg target. *Conform. Geom. Dyn.*, 18:119–156, 2014.

[12] Claude Dellacherie. *Ensembles analytiques, capacités, mesures de Hausdorff*. Lecture Notes in Mathematics, Vol. 295. Springer-Verlag, Berlin-New York, 1972.

[13] Szpilrajn E. La dimension et la mesure. *Fund. Math.*, 28:81–89, 1937.

[14] S. Eilenberg. On ϕ measures. *Annales de la Société Polonaise de Mathemmatique*, 17:252–253, 1938.

[15] Samuel Eilenberg and O. G. Harrold, Jr. Continua of finite linear measure. I. *Amer. J. Math.*, 65:137–146, 1943.

[16] Ryszard Engelking. *Theory of dimensions finite and infinite*, volume 10 of *Sigma Series in Pure Mathematics*. Heldermann Verlag, Lemgo, 1995.

[17] Behnam Esmayli and Piotr Hajłasz. Area and coarea formula for maps into metric spaces. (survey paper in preparation).

[18] Behnam Esmayli and Piotr Hajłasz. Lipschitz mappings, metric differentiability, and factorization through metric trees. (In preparation).

[19] Behnam Esmayli and Piotr Hajłasz. The coarea inequality. *Ann. Acad. Sci. Fenn. Math (to appear) arXiv:2006.00419v2*, 2020.

[20] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition, 2015.

[21] Herbert Federer. Some integralgeometric theorems. *Trans. Amer. Math. Soc.*, 77:238–261, 1954.

[22] Herbert Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[23] Piotr Hajł asz. Sobolev spaces on metric-measure spaces. In *Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002)*, volume 338 of *Contemp. Math.*, pages 173–218. Amer. Math. Soc., Providence, RI, 2003.

[24] Piotr Hajł asz, Mikhail V. Korobkov, and Jan Kristensen. A bridge between Dubovitskiĭ-Federer theorems and the coarea formula. *J. Funct. Anal.*, 272(3):1265–1295, 2017.

[25] Piotr Hajł asz and Soheil Malekzadeh. On conditions for unrectifiability of a metric space. *Anal. Geom. Metr. Spaces*, 3(1):1–14, 2015.

[26] Piotr Hajłasz, Soheil Malekzadeh, and Scott Zimmerman. Weak BLD mappings and Hausdorff measure. *Nonlinear Anal.*, 177(part B):524–531, 2018.

[27] Piotr Hajłasz and Scott Zimmerman. An implicit function theorem for Lipschitz mappings into metric space. *Indiana Univ. Math. J.*, 69(1):205–228, 2020.

[28] Juha Heinonen. *Lectures on analysis on metric spaces.* Universitext. Springer-Verlag, New York, 2001.

[29] J. D. Howroyd. On dimension and on the existence of sets of finite positive Hausdorff measure. *Proc. London Math. Soc. (3)*, 70(3):581–604, 1995.

[30] Witold Hurewicz and Henry Wallman. *Dimension Theory.* Princeton Mathematical Series, vol. 4. Princeton University Press, Princeton, N. J., 1941.

[31] Maria Karmanova. Rectifiable sets and coarea formula for metric-valued mappings. *J. Funct. Anal.*, 254(5):1410–1447, 2008.

[32] J. D. Kelly. A method for constructing measures appropriate for the study of Cartesian products. *Proc. London Math. Soc. (3)*, 26:521–546, 1973.

[33] J. D. Kelly. The increasing sets lemma, and the approximation of analytic sets from within by compact sets, for the measures generated by method III. *J. London Math. Soc. (2)*, 8:29–43, 1974.

[34] B. Kirchheim. Rectifiable metric spaces: local structure and regularity of the Hausdorff measure. *Proc. Amer. Math. Soc.*, 121(1):113–123, 1994.

[35] Steven G. Krantz and Harold R. Parks. *Geometric integration theory*. Cornerstones. Birkhäuser Boston, Inc., Boston, MA, 2008.

[36] Alexander Lytchak and Stefan Wenger. Intrinsic structure of minimal discs in metric spaces. *Geom. Topol.*, 22(1):591–644, 2018.

[37] Jan Malý. Coarea integration in metric spaces. *NAFSA 7—Nonlinear analysis, function spaces and applications. Vol. 7*, pages 148–192, 2003.

[38] Mathoverflow. Bounding an "integral" from below by the hausdorff measure of the domain. https://mathoverflow.net/q/355214/121665.

[39] Mathoverflow. Unknown work of nöbeling on topological/hausdorff dimension. https://mathoverflow.net/q/360384/121665.

[40] P. Matilla. Personal communications.

[41] Pertti Mattila. *Geometry of sets and measures in Euclidean spaces*, volume 44 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995. Fractals and rectifiability.

[42] Karl Menger. *Ergebnisse eines mathematischen Kolloquiums*. Springer-Verlag, Vienna, 1998. With contributions by J. W. Dawson, Jr., R. Engelking and W. Hildenbrand, a foreword by G. Debreu and an afterword in English by F. Alt, Edited by E. Dierker and K. Sigmund.

[43] G. Nöbeling. Hausdorffsche und mengentheoretische dimension. *Ergebnisse math. Kolloquium Wien*, 3:24–25, 1931.

[44] Anton Petrunin and Stephan Stadler. Metric-minimizing surfaces revisited. *Geom. Topol.*, 23(6):3111–3139, 2019.

[45] L. Reichel. The coarea formula for metric space valued maps. Ph.D. thesis, ETH Zurich, 2009.

[46] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[47] Leon Simon. *Lectures on geometric measure theory*, volume 3 of *Proceedings of the Centre for Mathematical Analysis, Australian National University*. Australian National University, Centre for Mathematical Analysis, Canberra, 1983.

[48] Stefan Wenger. Characterizations of metric trees and Gromov hyperbolic spaces. *Math. Res. Lett.*, 15(5):1017–1026, 2008.

[49] Stefan Wenger and Robert Young. Lipschitz extensions into jet space Carnot groups. *Math. Res. Lett.*, 17(6):1137–1149, 2010.

[50] Stefan Wenger and Robert Young. Lipschitz homotopy groups of the Heisenberg groups. *Geom. Funct. Anal.*, 24(1):387–402, 2014.